The Cognitive Reality of Prolongational Structures in Tonal Music.

Martínez, Isabel Cecilia.

THE COGNITIVE REALITY OF PROLONGATIONAL STRUCTURES IN TONAL MUSIC

BY

ISABEL CECILIA MARTINEZ

A thesis submitted in partial fulfilment of the requirements for the degree of PhD

SCHOOL OF EDUCATION

ROEHAMPTON UNIVERSITY

UNIVERSITY OF SURREY

2007

for Ana Laura, Juliana and Emilio

ABSTRACT

This thesis investigates the psychological implications of *prolongation, a structural phenomenon of tonal music*, which is described in the musicological literature as an elaborative process in which some pitch events - such as chords and notes - remain *as if* they were sounding even though they are not physically present. In spite of its theoretical value as an analytical device, the question of prolongation as experienced still remains. Its cognitive scope was thoroughly explored in two groups of experiments: in part 1 prolongation was hypothesized as a constituent organization, in which the linear continuity of the voice-leading as it unfolds is parsed into syntactical units with beginnings and endings. The listener's capacity to identify prolongational boundaries was tested under experimental conditions that explored the moment-to-moment sensitivity to prolongation in music-attending tasks. A clear ontology of prolongation as a constituent percept, at foreground reductional levels of the underlying structure, emerged unequivocally from the experimental results. However, this research did not fully explain an imaginative component that, according to Schenkerian theory, is present in the concept of prolongation. Alternative views of human cognition, related to the study of embodied knowledge and metaphorical thinking, were pursued in order to answer this question. In part 2 it was hypothesised that *prolongation would be experienced as* a *structural metaphor*. The interrupted structure, an archetypal organization of tonal music, was investigated on the assumption that this underlying configuration, interacting with cognition, primes in music perception the activation of *an image-schematic structure - involving force*. By means of a cross-domain mapping process, the listener projects that *image-schematic structure* onto the sonic organization of the piece, *understanding* the interrupted structure as a sonic

unfolding of the force schema. The results confirmed the hypothesis: prolongation has a

relevant status as *imagined cognition*. Structural metaphors operate as *idealized models of*

cognitive processing that listeners activate during their experience of music.

TABLE OF CONTENTS

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VII

EXPERIMENTAL STUDIES PART2: CROSS-DOMAIN MAPPING AND THE

CHAPTER VIII

ACKNOWLEDGEMENTS

CHAPTER I

THE CONCEPT OF HIERARCHY IN MUSIC THEORY

I.1 Modelling hierarchy: theoretical models of underlying structure

During the twentieth century, there has been broad agreement in the field of music theory about the idea that music structure is hierarchical. Based on that belief, different conceptions of hierarchy have been developed at different times, and hierarchy has become a relevant topic in academic discourse about music.

In general, hierarchical models aim at describing the structure and set up of musical pieces in terms of an ordered organization where constituent events are weighted differently, according to their relative structural importance within the sonic stream of the musical composition. The idea of hierarchy was informed during the second half of the twentieth century by developments in the fields of cognitive science and linguistics. Consequently, in some music-theoretical approaches, hierarchy is described in a way that is similar in all respects to descriptions provided by models of human cognition. This influence is not free of epistemological problems, given that music, as an artistic domain, might or might not be granted similar status to the one endorsed by the scientific and/or the linguistic domain. Therefore, direct applications of theoretical developments transferred from the fields of cognitive science and linguistics to the domain of art should be considered carefully.

Questions about cultural and/or stylistic vs. a specific musical piece's constraints, psychological vs. theoretical statuses of reality, and/or uniqueness vs. multiplicity of

attributions of structural meaning, arise when the concept of hierarchy is applied to music (Cohn and Dempster, 1992). In consequence, it is necessary to carry out more investigation in order to resolve some of these issues.

In the present section, the status of theoretical descriptions of music hierarchy will be explored in order to contribute to understanding the scope of its cognitive reality. Hierarchical structures have been regarded in music theory as entities that depict relationships of dominance-subordination between musical events. In hierarchical organizations, processes at superordinate levels - where musical information is represented in a general manner- monitor processes at subordinate levels, where data are represented in a more detailed manner.

Some hierarchical models that emerge from the canonical conception of contemporary music theory are called *models of underlying structure.* They present a view of musical compositions in terms of an underlying simplicity that unifies the diversity of the musical surface, and, as a consequence, endows the piece with coherence and integrity in order to achieve stability. Musical unity, according to this view, is not exposed in the complexity of the musical surface, but hidden in the underlying simplicity of the musical work. Unity, coherence and integrity are thus attributes that characterize the masterpieces of tonal music (Cohn & Dempster, 1992).

Proceeding in this way, models of underlying structure "attempt to capture a level of experience at which a work is grasped as a single pattern or unitary structure rather than as a concatenation of patterns or composite of parts" (Benjamin 1982, p. 28). What counts as important in models of underlying structure is that they express relationships between sets of events, rather than single events. In these models, hierarchical relations take the form of a one-to-many correspondence.

One of the pioneers of considering hierarchy in music - and the first who understood it as a one-to-many correspondence - was the German musicologist Heinrich Schenker ([1906] - 1990); ([1910 - 1922] - 1987); ([1935] - 1979), who developed a model of underlying musical structure. At the time Schenker's theory was conceived, the concept of hierarchy was not new. It had implicitly been developed in the rules of strict counterpoint. For example, in the context of species counterpoint the cantus firmus acted as an organizing principle that provided coherence to a given musical passage, and the cantus tones were literally heard throughout, determining time-spans in the piece. It is an assumption of Schenkerian theory that the characteristic durational quality of the tones of the cantus firmus, by means of which they maintain its sonic presence through time, is transferred to the elaboration of tonal music in the so called *free composition style* (Schenker ([1935] - 1979), endowing the more structural tones with the quality of being somehow present throughout extended time-spans, forming an imaginary extended melodic line, even though they have no physical existence. This issue will be discussed later in this chapter.

Schenker's view is based on musicological arguments about the primacy of the horizontal dimension in music. The temporal development of a musical work is understood as the *unfolding of the linear component* of the piece. Relationships between musical events all along the underlying hierarchy are ruled by basic components of tonal organization. The idea of an organic force that provides coherence to the musical piece is at the core of his conception of the underlying musical structure. Even though Schenker's writings contain countless assumptions about the *psychological* status of his analytical developments, his model was not explicitly intended to be a model of musical competence in a Chomskian sense, but a model of musical aesthetics. It was not until recently that some of the theoretical constructs of Schenker's hierarchical theory were formalized in the Generative

Theory of Tonal Music by Lerdahl and Jackendoff (1983) as a model of ideal musical competence.

A hierarchical analysis may reflect the form of a musical piece to a greater or lesser degree. That is to say, the underlying structure of a composition may express, at a given moment, a more or less remote representation of the musical surface of the piece, according to the level of abstraction at which the analysis is performed. As abstraction progresses from a literal rendering of the surface to deeper levels of its underlying structure, the number of events contained in the model's representation decreases. Thus, according to Benjamin (1982), it is possible to say that if two models represent the underlying structure of the same piece and both of them positively capture the same components, the model that is more abstract is the one that contains fewer events.

However, there is no theoretical agreement about the nature of underlying hierarchical structures. Benjamin considers the manner in which hierarchical representations are derived from musical pieces and identifies several criteria by which a given underlying representation may be more or less remote as an abstraction of the piece it represents. Interesting for our purposes is the extent to which events' representations are copies of events in the piece:

> "Proceeding in one way, one might listen to a passage within a work and determine that, during this passage, there occurs, as phenomenon of the piece, an event which seems most stable or focal and to which it seems best to relate all other events comprising the passage. One might then carry a copy of this event up into one's model of the passage by considering that copy to represent the passage as a whole at the level of the model. This procedure makes for a literal relationship between model and piece: the most important surface event is borrowed from the piece and becomes the underlying event of the model. Quite another relationship results when a different procedure is followed. This involves considering the underlying event to be somehow

manifest throughout the relevant passage […]. Here, the relationship between model and piece

may be described as symbolic; the single event of the model symbolizes the unity underlying

the multiplicity of events in the passage. It is perhaps worth emphasizing that in neither case

are model events the same as piece events; it is just that in one case they are copies while in the

other they are not" (Benjamin 1982, p. 32).

In the first procedure, that accounts for a literal relationship between model and piece,

events are extracted from the musical surface and *behave as literal copies that represent*

the piece. The second procedure, on the other hand, describes a *symbolic* relationship

between model and piece, in which events are endowed with a property to manifest as

abstract representatives that convey unity throughout the complexities of the piece.

Analogously, Cohn and Dempster (1992) proposed two types of *ontologically different*

hierarchies that contribute to the analysis of models of underlying structure: (i) the

underlying structure is a *representational hierarchy*, where a musical set of pitch events is

dominated by one of its constituent events, the latter being consequently promoted to a

superordinate level in the hierarchy. A representational event, therefore, is a concrete pitch

event with specific features of location and duration in the musical composition. In this

ontology, events at different levels of the hierarchy differ only in their structural

importance, being similar in type. (ii) the underlying structure is an *inclusional hierarchy*,

where abstract classes of events control sets of pitch events in the musical composition,

without being literally equivalent to any of them. In this ontology, the relationship of the

abstract representatives with the concrete pitch events is metaphorical and heuristic rather

than literal.

Whether they are representational or inclusional, these two alternative views of the

underlying structure, the one that builds up the underlying hierarchy based on literal

representations of the musical surface and the other that conceives underlying

5

representations as symbolic or metaphorical, account nevertheless for unifying principles that govern the internal constitution of a musical composition. In so doing, they "reduce the many and much to one and little" (Cohn & Dempster 1992, p.156). As far as the apprehension of musical structure is concerned, these ontologies have implications for the study of underlying musical structure, and prompt assumptions about the use of potentially different cognitive processes in the listener's experience of musical pieces.

However, the key question is whether these theoretical principles of music hierarchy are simply constructs that account for an ability to devise analytical systems or if they have a psychological reality. The answer to this question, concerning the concept of prolongation as a central principle in models of underlying structure, is central to the present study.

I.2 The concept of prolongation

I.2.1 Prolongation as a component of underlying musical structure

Prolongation is a structural phenomenon described in the literature of music theories of underlying structure (Schenker [1935] - 1979, Lerdahl & Jackendoff, 1983; Salzer [1962] - 1982; Salzer & Schachter, 1969) as an elaborative process in which some pitch events - such as chords and notes - remain (*as if* they were sounding) within the musical stream of events, even though they are not physically present.

A prolongational analysis of a group of tones in a musical passage would usually be performed by assigning the most stable tone a structural role, and the other tones an elaborative role relative to the former tone. It is a strong assumption that the governing sonority of the stable tone yields a trace that manifests itself throughout the subsidiary tones that elaborate it, until a subsequent melodically stable tone appears. In prolongational

models, then, prolonging events somehow evidence the presence of the structural, prolonged event by being subservient to it as if it were revealed all through the time-span of the musical passage. Thus, it seems as if prolonging events behave as a *metaphorical extension* of the prolonged event (Benjamin, 1982).

Therefore, by means of the elaborative processes prompted by prolongation, the unfolding of this kind of one-to-many relationship connects non-immediate events at a more remote level of the underlying hierarchy. It is assumed that such a relationship is replicated throughout the underlying structure of the musical composition, accounting for an organization formed by a number of hierarchical levels: this is shown in Figure I.1.

A reductional analysis of this, Schenkerian, type, of the prolongational arrangement of a musical piece will reveal the fundamental scale-step structure that underlies the music. It is an assumption of Schenker's prolongational theory that the linear elaborations that result from such an arrangement establish strong connections between the different scale-steps.

Figure I.1: (a) Beginning of Beethoven's Piano Sonata Op. 110, I (bars 1-5); (b) foreground reduction of the fragment; (c) middleground reduction of the fragment. (Extracted from Cadwallader and Gagné 1998, pp.136-137).

Therefore, by means of prolongation, the underlying structure moves forward from one scale-step to another and the strong connections between those fundamental tones bring tonal coherence to the whole music organization. We can see in Figure I.2 an example of this.

Figure I.2: (a) Beginning of Haydn's Piano Sonata Hob. XVI/35, I, (bars 1-8); (b) analytical reduction showing in the upper line a descent step-motion from $\hat{5}$ to $\hat{1}$ and in the bass a harmonic progression unfolding the structural bass-line pattern of I-II^6-V-I; (c) final level of reduction. (Extracted from Cadwallader and Gagné 1998, pp.112-113 and 117).

According to this view, prolongational processes depict a landscape in which the presence of the fundamental melodic-harmonic tonal components is extended; in that landscape, the underlying structure seems to function as an organic force, a kind of abstract entity that brings to the whole musical composition a simplifying unity.

The models of underlying structure present a particular view of musical motion over time. Benjamin (1982) makes a distinction between *prolongational* and *progressional* models of underlying structure, according to the path that the model follows. In prolongational models, the unfolding of the musical piece fixes motion to the non-temporal structural configuration of the fundamental structure, and the complexities of the voice-leading as they unfold are regulated by the characteristic shape of its underlying path. The kind of change produced by the elaborative processes of prolongation qualifies metaphorically as movement in a continuous line between points that are structurally and significantly related. By means of prolongation, the pitch-class structure of the piece (its scale-degree organization) moves from one structural place to another. The path is essentially linear and motion is phenomenally interpreted as being continuously within some fundamental structural pattern such as, according to Schenker, for example, a triad.

Progressional models present a rather different picture of the unfolding of musical structure. Such models are prolongational only at a local level, but at a global level the whole design of the musical composition depicts relationships between chains of musical segments, in such a way that motion from one segment to another generates angles of deviation between successive lines, rather than motion throughout a continuous line. What is important in these models is their progression from one line of motion to another.

Progressional models had been previously developed in the works of Rameau and Riemann, but were not intended to be informative about any specific underlying ontology.

In prolongational models the quality of motion generates a path that is conceived as something essentially and continuously linear, an ongoing musical flow that evolves within a fundamental structural pattern, or, in other words, between points that are significantly related by step movement. As a consequence, the unfolding of such continuous linearity exhibits a quality of a unifying wholeness. In these models "…piece events which follow the first structural event but precede the second, manifest the influence of the former as a presence which is somehow felt throughout their collective time-span. Prolonging events, therefore, may be thought of as extending a preceding structural event by continually clinging to some concrete manifestation of the latter…" (Benjamin 1982, p. 48).

By contrast, in progressional models, the overall design presents a course of action characterized by frequent departures of the initial line of motion, sometimes contradicting it instead of being permanently subject to it. In progressional models each intermediate event obtains its function as a result of its connection with the immediately surrounding events, that is to say, the event is merely a link in a chain of events that has the initial and terminal events of the chain as the most structurally salient. However, it is not the function of the intermediate events to prolong either the initial or the terminal event. A progressional model is highlighted when the modulational plan of a tonal piece is analysed (Benjamin, 1982).

Even though both approaches have frequently been presented in the literature of music theory as somewhat contradictory views, and what's more, their priority as generative principles of tonal structures has been questioned (see Horton, 2003), prolongation and

progression may nevertheless be interpreted as two interrelated structural phenomena that provide coherence and direction to the whole process of linear elaboration.

The underlying structure that results from the linear processes prompted by prolongation represents, according to Schenker, a sort of internal voice, on which tonal coherence is based. He calls it the *underlying voice-leading.* The organization conveyed by the underlying voice-leading is the result of a number of elaborative processes in which melodic and harmonic components are combined. In this organization, the inner voices of the fundamental structure are subordinated to the outer voices, that is to say, to the fundamental line and the bass arpeggiation (Schenker [1935] - 1979, p. 17). It is in the middleground (see below section I.2.2, p. 13) that we find the most typical voice-leading transformations that elaborate the fundamental structure. This is the reason why some current theorists of tonal music (Benjamin, 1981; Komar, 1992) argue that traditional harmony should be reconsidered, including in the analysis of the chordal theory the perspective provided by the theory of lines, given that the roles played by both intervals and dissonance in determining progressions are more the result of the composing-out of lines and pitch-class lines than of chords. That is to say, "tonal harmonic progressions are counterpoints of four pitch-class voices, motion in one of which is determined by motion in one or more of the others. Therefore, it is the complex of lines in a progression which determines its identity, and not the chords of which it is comprised" (Benjamin 1982, p. 40).

Prolongation, as an attribute of the underlying voice-leading, is characterized in post-Schenkerian approaches (Salzer [1962]-1982; Salzer & Schachter, 1969; Forte & Gilbert [1982] - 1992; Cadwallader & Gagné, 1998; Schachter, 1998) as a comprehensive expression that conveys different composing-out ideas, such as "the elaboration,

development, manipulation and transformation of underlying principles" (Salzer &

Schachter, 1969, p. xix).

Since prolongation is a concept that is copiously employed in music analysis, and given

that its value as an essential analytical tool for the treatment of hierarchical structure in

music is usually taken for granted in different theoretical sources, its potential

psychological implications deserve investigation. As stated above, this is the main purpose

of the present study.

I.2.2 Prolongation in three hierarchical theories of music

I.2.2.1 The model of underlying structure by H. Schenker

H. Schenker ([1925] - 1994; [1926] - 1996; [1910 - 1922] - 1987; [1935] - 1979) presented

a model of underlying structure and developed the idea of organic coherence in music.

Organic coherence arises from the hierarchical organization of musical pieces, and is based

on musicological arguments related to the unfolding of the tonic triad. According to

Schenker, the composing-out of a piece of music is governed by the principles of

counterpoint and voice-leading, which give pieces musical direction. In Schenker's

reductional analysis, the underlying musical structure is unfolded across three hierarchical

levels: the background, the middleground and the foreground (*Hintergrund*, *Mittelgrund*

and *Vordergrund*, according to Schenker [1935] - 1979). These three levels elaborate the

structure of the composition from its deepest state to the musical surface. The background

holds the fundamental structure (*Ursatz*) that underlies the whole composing-out process.

The fundamental structure is the result of the counterpoint between the fundamental line, in

the upper register, and the bass line, in the lower register. The fundamental line (*Urlinie*) is

the melodic component of the fundamental structure; it unfolds a sequence of descending

scale-steps (*Stufen)* that can fill-in the interval between the third and the tonic (3, 2, 1), the fifth and the tonic (5, 4, 3, 2, 1) or the eighth and the tonic (8, 7, 6, 5, 4, 3, 2, 1) scale-steps. The bass voice is the harmonic component of the fundamental structure; proceeding by disjoint melodic movement (*Bassbrechung*) it unfolds the interval of an ascending fifth, between the tonic and the dominant scale-steps, and returns downwards to the tonic (I-V-I). The fundamental line and the bass arpeggiation combine with each other to form a pattern that is intended as an *archetypal unity* (see Schenker, [1921-1924] 2004 - 2005). In Figure I.3 the fundamental structure is shown.

Figure I.3: (a) The Fundamental Structure with the upper fundamental line and the bass arpegiation; (b) Obligatory register of the fundamental line (changes of register occur at middleground and foreground levels. (Extracted from Schenker [1935]-1979), figs. 1 and 4).

Schenker was the first who proposed the idea of prolongation (*Auskomponierung*) in his theory of underlying structure. According to him, prolongation begins in the *Ursatz.* Processes of transformation of the tones of the fundamental structure operate from the background to the foreground, which "horizontalizes" the background.

Athough the Schenkerian approach appears to assign primacy to the linear aspects of music unfolding, prolongation, in Schenker's view, is the result of the combination of harmonic

and melodic dimensions. As pointed out above, the contrapuntal origin of the foreground level, its derivation from voice-leading procedures that horizontalize an underlying state, and the subsequent re-applications of the one-to-many relationship between pitch events at each of the three hierarchical levels of the embedded musical structure (the background, the middleground and the foreground) are the corner-stones of his theory of the underlying structure.

When a Schenkerian analysis is performed, the musical composition is scanned in order to unveil the application of a number of prolongational techniques such as the initial ascent, the linear progressions, the motion into and from an inner voice, the reaching over, the transfer of register and the substitution, among others, responsible for the composing-out and continuous disclosure of the background, middleground and foreground levels and, in the end, the fulfilment of the entire structure of the piece (see Cadwallader & Gagné, 1998 for a full account of the prolongational techniques).

The concept of organic coherence, in which prolongation is embedded, is an idea coined in the context of Schenker's philosophical foundations (Blasius, 1996). The metaphor of organicism and the biological terms associated with it govern his theory of the underlying structure and are present in the language Schenker uses to describe the processes of prolongation that take place when the underlying structure is unfolded. For instance, he characterizes the fundamental structure as a living organism, enlightened by a vital natural power at play in the musical work. The fundamental line is conceived as a path in which directed motion is propelled by a sort of life-impulse that puts different forces in operation. Traversal of this path is accomplished in a downward, straight-forward manner throughout the different levels of the underlying hierarchy. He conceives the linear motion as being forced by the necessity of reaching a goal, the tonic. "...we feel by nature that the

fundamental line must lead downward until it reaches 1 and that the bass must fall back to the fundamental" (Schenker [1935]-1979, p. 13). Tonal coherence, finally, is the result of the inevitable continuity of motion in the same direction, striving toward the intended goal until the course of the unfolding procedure is completed.

At the core of the Schenkerian conception it is the idea that prolongation can be ascribed the status of an implied entity in which discrete events within the musical composition are mentally connected. Thus, the effect of prolongation is that of a *metaphorical extension*, in the sense that somehow the tone that is prolonged *is* present throughout the prolonging events of a given passage. The structural dominance of the prolonged event over the prolonging events is an emergent property of the above-mentioned arguments of organic coherence, underscoring the ideas of propelling force and continuous motion throughout the fundamental line of the underlying structure. Some psychoacoustical arguments that provide strong explanations of the sonic architecture of tonal music are consistent with the notions of structural dominance of the prolonged event (see Chapter II section II.2.1 for an explanation of cognitive principles of tonal organization). According to them, the tonal quality of a prolonged tone will resist sonic assimilation by subsequent prolonging tones within the stream of events of a musical passage. This ontology of prolongation, then, assumes that the prolonged event *sounds as if* it were a kind of virtual pedal (Benjamin, 1982).

The idea of metaphorical extension is also congruent with another notion coined in Schenkerian tradition related to the operation of a latent cognitive mechanism in the experience of prolongation called *mental retention*, referred to as a form of prolongation in which a tone is mentally represented and *kept in mind* until the next tone of a structural upper voice becomes active (see Cadwallader and Gagné 1998, p. 401, note 21).

Schenker's view of the underlying structure implies the idea of an *inclusional hierarchy* (Cohn and Dempster, 1992). The embedded tonal organization that he proposes is unveiled as the musical work unfolds along a continuous time-line in which prolongational events are not understood merely as concatenations of unbroken units but, as entities deep-rooted in layered nestings of increasingly longer temporal spans. Schenker's conception of tonal coherence results from the deployment of these mutually embedded underlying levels that provide containment of and fulfilment to the configuration of the whole composition.

As stated before, counterpoint and voice-leading are two building blocks in Schenkerian theory. Schenker's conception of underlying voice-leading, however, includes a prolongational harmonic view. In his view, scale-steps serve more the projection of the horizontal rather than the delimitation of tonal territories - as occurs in progressional harmonic approaches - because what counts as important is that a scale-step's succession is a necessary consequence of the prolongation of the tonic sonority that underlies it, throughout a musical composition (Benjamin, 1982). The harmonic dimension, then, is best understood in his theory not just as the chordal manifestation of the vertical domain of the musical piece, but as the capacity of scale-degrees to *move* - through prolongation - toward the achievement of the tonic.

Schenker's conception of the relationship between harmony and counterpoint in the elaboration of musical structure leads him to assert, "…the harmonic degrees are inextricably bound up with counterpoint" (Schenker [1935]-1979, p. 35). The bass has a primary importance in that it provides unity to the fundamental chord. Scale-degrees are present in the bass, where they unfold the fundamental structure. Thanks to their strength, scale-degrees play a crucial role in guaranteeing tonal coherence. As a consequence, their control is found throughout the middleground and the foreground.

"Throughout the prolongational levels it is sometimes the horizontal which determines the particular course and meaning of the vertical, sometimes the vertical which by its own voice-leading dictates the horizontal. So, in one instance we can speak of the horizontalization of the vertical, in another instance of the verticalization of the horizontal. It would be erroneous therefore, to read all degrees in the foreground without discriminating between them, as though they were all of equal significance and origin. Rather one must make the following distinction: between those harmonies that, in a particular way, serve particular diminutions close to the foreground, and those harmonies which, in their origins, express strong relationships in the levels close to the background" (Schenker [1935]-1979, pp. 111-112).

According to Schenker, the logic of voice-leading ensures onwards motion to a goal. Consequently, individual harmonies that emerge from the progression of the various voices are also forced to move forward, and those transitory harmonies that appear mainly at foreground level are the consequence of the inner necessities of voice-leading. Conventional notions of harmonic progression do not apply consistently to the analysis of the surfaces of many tonal works, because there are *piece states* (Benjamin, 1982) which, according to Schenker, are much more efficiently described in contrapuntal terms. In those cases it is the underlying voice-leading that governs harmonic deployment and not the other way round. For example, one of the core concepts found in Schenkerian theory is that of *linear progressions* (Schenker [1935]-1979, p. 9). They represent the melodic stepwise motions that span different harmonic intervals, thus horizontalizing the vertical dimension, whether within a single harmony or between different harmonies. Schenker considers them to be anchored in polyphony and proposes to understand them in contrapuntal terms. Moreover, he attributes a primary role to them as a means to provide coherence to the unfolding of the underlying structure of musical compositions.

Therefore, Schenker's proposed harmonic linearity should not be understood as the result of processes that take place over successive harmonic states in which scale-degrees of an initial harmony are successively displaced by adjacent scale-degrees, which, although foreign to that harmony, have equal importance in the succession of events. On the contrary, the implied linear-harmonic organization composes-out a peculiar arrangement of events in which some predominate over others in order to form, at a more remote hierarchical level, a fundamental harmonic progression. In order to fill-in, through prolongational processes, the tonal space between an initial tonic, a consecutive dominant and a conclusive tonic in the bass line at the level of the fundamental structure, all the intermediate available harmonies may be included (Deliège, 1984). But the elaborative processes that unfold the intervals between the scale-degrees of the fundamental progression using intermediate harmonies are governed by the shape of the fundamental structure, in order to preserve, at the deepest level, its basic configuration both in the bass and the fundamental line.

Concerning the temporal scope in the unfolding of the harmonic structure, surface harmonic progressions do not extend in time beyond the temporal spans of their individual event's states. In contrast, harmonic prolongations at deeper levels account for the control of longer time-spans. The basic pattern underlying this organization is a non-modulating diatonic structure, and the components of its 'architectural building' are the secondary processes that emerge from the prolongations, transforming the basic structure in such a way that the shape of its archetypal pattern is preserved.

One of the keystones of the concept of prolongation is that it conveys the idea of establishing distant connections between non-adjacent tones in the sequence of events of the musical piece. By proceeding from the background to the foreground, the generative

19

elaboration progressively transforms the fundamental structure into longer streams of events, decreasing the time-span between the non-adjacent events until the musical surface is reached. The opposite process is the reduction of events at the surface level into progressively deeper levels of the underlying structure, where increasingly distant connections between pitch events are established, and non-structural pitch components are progressively eliminated until the fundamental structure is reached. Concerning the way in which reductional processes operate, C. Schachter posits:

> "What this means is that one can never hope to arrive at a correct view of the background by simply making a "reduction" of the foreground, for example, by eliminating dissonances, chromatics, or non-tonic notes. Without some sense of the background, one can´t begin to understand the foreground; it might be precisely those dissonant or chromatic elements a reduction would eliminate that form the "background" of a passage. But if one needs to understand the background to make sense of the foreground, one also needs to understand the foreground to make sense of the background – a seemingly hopeless impasse. Actually, it's a heuristic problem that confronts people all the time and in areas far removed from musical analysis: one can grasp neither the part without the whole nor the whole without the part" (Schachter 1981, p. 132).

Implicit in the analysis of the underlying structure is the idea that reductional procedures are better understood in heuristic terms (procedures that operate as vehicles for solving problems by discovery, reducing the range of possible answers to a question) rather than in algorithmic ones (procedures where the solution to a particular problem is guaranteed by the mechanical fulfilment of a finite number of steps). Moreover, the relationship between the background, the middleground and the foreground is dialectic, in the sense that one is permanently prompted to *traverse the path* between the different hierarchical levels, and to make comparisons between different events and the entire whole in order to draw a

convincing picture of the landscape of the piece, as long as the analysis of the unfolding structure progresses.

The copious literature on Schenkerian analysis developed during the past half century in the field of Western musicology, both European and American, attests to the further developments of the theory that Schenker's followers accomplished (see Salzer & Schachter 1969; Salzer [1962] - 1982; Forte & Gilbert 1982; Cadwallader & Gagné 1998, Schachter 1998, to mention a few).

I.2.2.2 The generative theory of tonal music by Lerdahl and Jackendoff

The generative theory of tonal music *(GTTM)* by Lerdahl & Jackendoff (1983) formalises a structural description of music's organisation in terms of systems of rules based on principles inspired by those of generative linguistics.

In *GTTM,* listeners' familiarity with the musical idiom elicits the organization of sound sequences into coherent structures, the latter resulting from the operation of psychological principles that are unconsciously applied during the aural analysis of the sonic signal of the musical piece. Structure is therefore an inference, and the model's formalization intends to account for aspects of the *heard* organization of the composition. Four types of structural system are derived from the generative analysis of musical pieces: grouping structure, metrical structure, time-span reduction and prolongational reduction. Three principles - no overlapping, adjacency and recursivity - guarantee the well-formedness of those structures.

The idea of prolongational hierarchy proposed in *GTTM* results from the instantiation of a temporal reductional process, called *time-span reduction*, that occurs within the grouping and metric structures of the musical piece. The temporal span becomes a regulatory factor

of pitch stability and is used to determine the structural importance of an event, as is shown in Figure I.4.

Figure I.4: Time-span reduction of the first phrase of Bach's chorale "Christus, der ist mein Leben". It can be observed a close correspondence between the tree notation, that appears in the analysis instantiated in the constituent tree above, and the description - in musical notation – that appears in the analysis of the time-span reduction presented below. (Extracted from Lerdahl 2001, p. 11).

Lerdahl & Jackendoff tackle prolongation in a rather different way from Schenker; prolongation in *GTTM* is an attribute that emerges from the application of the prolongational reduction analytical technique. According to the theory, prolongation is viewed in terms of patterns of tension and relaxation, which depend on the *relative stability* between events. The basic hierarchic relationship between two events is that of an *elaboration*, in which, if an event is the elaboration of another event, the first one is structurally more important than the second one, and additionally more stable.

Prolongation is analysed according to three possible types of connection between two successive events, taking into account their rhythmic, melodic and/or harmonic dimensions and the tensing/relaxing functions that result from their combination. This is indicated by the direction of the branching in the tree diagrams shown in Figure I.5 below:
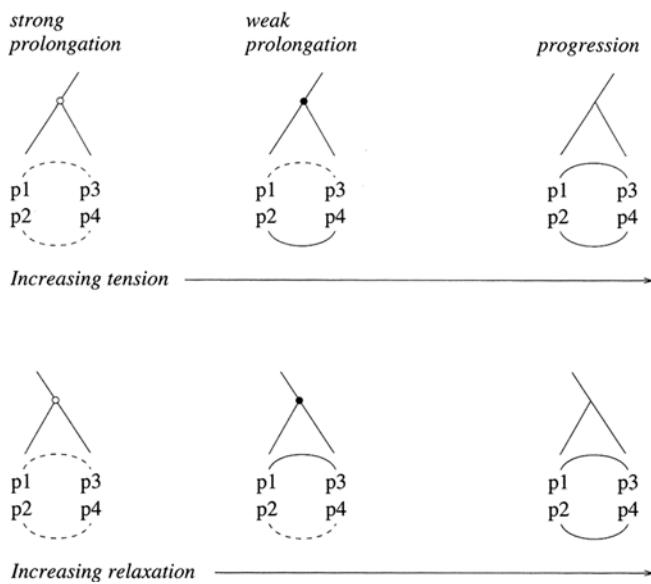
Figure I.5: Types of prolongational branchings. The tree notation projects patterns of tension and relaxation and the slur notation conveys linear connections. Right branching corresponds to increasing tension while left branching shows increasing relaxation. Dashed slurs correspond to strong prolongations, a combination of dashed and solid slurs corresponds to weak prolongations and solid slurs correspond to progressions. (Extracted from Lerdahl 2001, p .15).

23

(i) *strong prolongation*, where the notes from the melody and the harmonic fundamentals are identical (in this case, tension and relaxation are exclusively due to the incidence of rhythmic factors); (ii) *weak prolongation*, in which the harmonic fundamentals of the two events are identical but one of them is in a less consonant position - with regards to the bass or the other melodic note -; and (iii) *progression*, which occurs when the harmonic roots of the two events are different.

What makes one event more prolongational than other in *GTTM*? The theory posits that pitch stability is necessary but not sufficient as a single criterion to determine the degree of stability of a given event, and highlights the importance of the rhythmic dimension and its primacy in the perception of prolongational relationships: according to this view, events that are not rhythmically important (weak-accented) are heard as less important prolongationally, while strong-accented events are interpreted as structurally important, no matter the pitch. As will be shown later, this rhythmic constraint - relative to the listener's experience of the prolongational structure - does not necessarily apply (see Chapter IV below).

This view of pitch stability is rather different from the one depicted in Schenkerian theory. In *GTTM* the temporal span becomes a regulatory factor that is used to assess the prolongational importance of an event. The close dependence of the prolongational reduction on the time-span is one of the cornerstones of the theory.

The method used by Lerdahl & Jackendoff to perform the reductional analysis also differs from the one used by Schenker. It operates recursively from the local to the global level of the structure, reducing the temporal span from the musical surface to the *Basic Form* (the term used by Lerdahl for Schenker's fundamental structure). Once the time-span reduction

has been realized recursively, the reduced structure obtained is then processed from the global to the local level, in order to obtain the prolongational reduction (Lerdahl, 2001).

The basic form is the abstraction of the progression I-V-I, which in *GTTM* is described as a common reductional state that accounts for the whole structural course, from the beginning to the final cadence, of the musical piece. Lerdahl (2001) points that, even though *GTTM's* basic form superficially resembles Schenker's *Ursatz,* it is not conceived in the same way. Instead of being assumed –as in Schenker's theory - as an *a priori* structure, the basic form is generated as the result of a number of repeatedly concurrent factors that provide stability to the structure of tonal pieces.

> "This basic form occurs as often as it does because it realizes a number of converging stability factors that are simple and frequently available at or near the musical surface: prolongation of the tonic, root-position harmonies, the standard cadence, and diatonic downward stepwise melodic motion that resolves on 1" (Lerdahl 2001, p. 25).

Prolongation in *GTTM* is the result of understanding musical structure in terms of a *representational hierarchy*. In representational hierarchies, higher-level and terminal events differ in their relative structural importance but not in kind (see section I.1 above, p.5). A constituent unit is dominated by one of its constituent events (i.e. a literal pitch event). As a result of the reductional analysis, that event is selected from the stream of events and promoted one level upward in the hierarchy. In order to perform the selection procedure and assign to each of the specific events their prolongational status, *GTTM* takes into account the external boundaries of the progressively reduced structures, and selects from those places the literal structural tones to be promoted. Thus, in Lerdahl & Jackendoff's hierarchies, selected events at a given hierarchical level are literal representatives of events at a different hierarchical level.

GTTM fails to provide a detailed, extensive description of prolongation, according to the principles of voice-leading, such as that proposed by Schenkerian theory. The shortcomings are related to difficulties in the application of generative analysis in the prolongational reduction of harmonic progressions, in order to formalise the voice-leading of the independent horizontal lines. Should such an application be pursued, it would be necessary to perform an independent overlapping analysis for each voice, thus complicating *GTTM's* proposed formalisation.

In later applications of *GTTM*'s principles (Lerdahl 1988, 1997, 2001), some aspects of the theory were revised. For example, the ontology of underlying events was reconsidered. Selected events became to be understood as abstract representatives (somehow resembling the Schenkerian view), generated through the operation of transformation, in which some events are mentally retained while others degrade in memory, according to the contextual conditions of the structural hierarchy in which those events operate. But the kind of abstraction proposed in the revised version of *GTTM* is different from the Schenkerian one. *GTTM*'s idea of abstraction as mental retention postulates that the specific content of an event degrades and generalizes in memory, depending on the contextual framework in which it is experienced. Consequently, at any prolongational level, only what is needed in that context is retained through a transformational operation. The transformation proposed by Lerdahl consists of the deletion of surface events. There are two types of abstractions produced by the transformation process: event abstraction, explained above, and temporal abstraction, which is related to the verticalization of events within the time-span reductional process. In temporal abstraction, those events that are located at a lower level in a weak metrical position, once abstracted, occupy a strong metrical position at the next

higher level. Nevertheless, a problematic issue is posited when this procedure is considered:

> "When to activate and how to specify the content of transformations [ ] is theoretically
> problematic.[ ] often it is not as clear [ ] whether an event is suppressed at the surface level
> and if so, what it should consist of at underlying levels. Abstracting event features at the level
> of the phrase [ ] may seem relatively straightforward, but what the nature of an event is at a
> global level [ ] is far from obvious" (Lerdahl 2001, p. 37).

In order to overcome *GTTM*'s shortcomings alluded to above, a more accurate description of the horizontal aspects of the prolongational structure is provided, but still strongly related to vertical, homophonic aspects of the musical piece. Lerdahl (2001) develops a theory of relative pitch attractions which he applies to the analysis of harmonic progressions. In this approach, while seeking stability, each of the pitches of a given chord is assumed to tend towards another tone that is considered an 'attractor' in a given context. With the intent of synthesizing certain advances in the analysis of the psychological implications of hierarchy, the description provided by Lerdahl's theory of melodic and harmonic attractions is informed by certain basic psychological principles of music cognition proposed elsewhere, such as those provided by the implication-realization model of Meyer (1973) and Narmour (1990; 1992), the melodic anchoring model of Bharucha (1984a; 1996) and the application of some Schenkerian principles to a theory of melodic forces by Larson (1997). (See chapter II for a description of relevant principles of music cognition).

I.2.2.3 The implication-realization model by L. Meyer- E. Narmour

The implication-realization model (*I-R*), based on the theory proposed by Meyer (1973), but developed and formalized later by Narmour (1977; 1983-1984; 1990; 1992), is another

music-theoretical framework that tackles the problem of structural hierarchy along with the concept of prolongation. Based on psychological assumptions about the relationship between music and basic cognitive mechanisms in action while a musical piece is experienced (as described, for example, by the Gestalt laws of common fate or good continuation) this approach focuses primarily on the analysis of the melodic component of compositions. Melodic surfaces are treated largely as note-to-note phenomena, in which the continuity of melodic relations is substantiated through the structural processing that operates at lower levels of musical organization.

The formalization proposed by *I-R* aims to track and reflect psychological features related to the fluctuations in arousal that occur as a consequence of the validation or invalidation of the moment-to-moment subconscious inferences made by a listener, concerning the way that a piece of music is likely to unfold. This kind of cognitive processing is founded in the differentiation between simple melodic patterns in terms of their implications for continuation. The implication of continuation ("process") and the implication of differentiation ("reversal") arise through the operation of Gestalt principles of similarity, proximity or common direction, when consecutive pitches lie in relative close proximity to one another. Two formal hypotheses - that similarity between elements or events implies continuing similarity (A+A implies A), and that differentiation implies continuing differentiation (A+B implies C) - produce a small number of patterns that function as "melodic archetypes". The five melodic archetypes Narmour identifies are as follows:

1. process or iteration (A+A, nonclosural); 2. reversal (A+B, closural); 3. registral return; 4. dyad (i.e. two-element groupings) and 5. monad (i.e., one-element groupings, closed or unclosed, where no generation of implication occurs). (Narmour 1990, p.45-5).

The reductional status of the *I-R* model implies prolongation, but its scope is restricted to the process of implication. Thus, in an implicative pitch pattern, that is, a pattern *that seems to require* a realization, prolongation is the result of maintaining the physical instantiation of the notes in the implicative pattern until realisation is reached. Prolongation, in this context, depends on the peculiarities of the musical surface, mainly on the implicative reduction of the individual constituent melodic lines.

Narmour applies the idea of "closure" to the analysis of those archetypal patterns that result from the implication-realization process. Closure is defined in terms of the relative strength of the melodic implication, brought about by its realization or its denial, the idea of process being nonclosural and the idea of reversal fundamentally closural (see above). Closure functions as the mechanism whereby *multiple levels* may be generated in the cognition of melody, and also determines the way in which patterns of melodic notes may be grouped. By means of closure, then, a structural hierarchy is generated from the musical surface, and the first higher level, the one immediately above the musical surface, typically consists of the notes located at the beginning and at the end of a given surface pattern. Hereafter, the principles by which patterns are obtained may be applied recursively, where each further application generates a higher-level description. Strong closure produces the "transformation" of notes (that is, their selection as events functioning at a structural level above that of the musical surface) and thus is the base for the reduction of melodic components of the musical surface. But, according to Narmour's theory, as far as structure is concerned, the nature of the internal unfolding of a melodic pattern, through the note-to-note processes of implication-realization, is as important as the way the pattern begins and ends. An example of the application of the implication-realization reductional analysis is shown in Figure I.6.
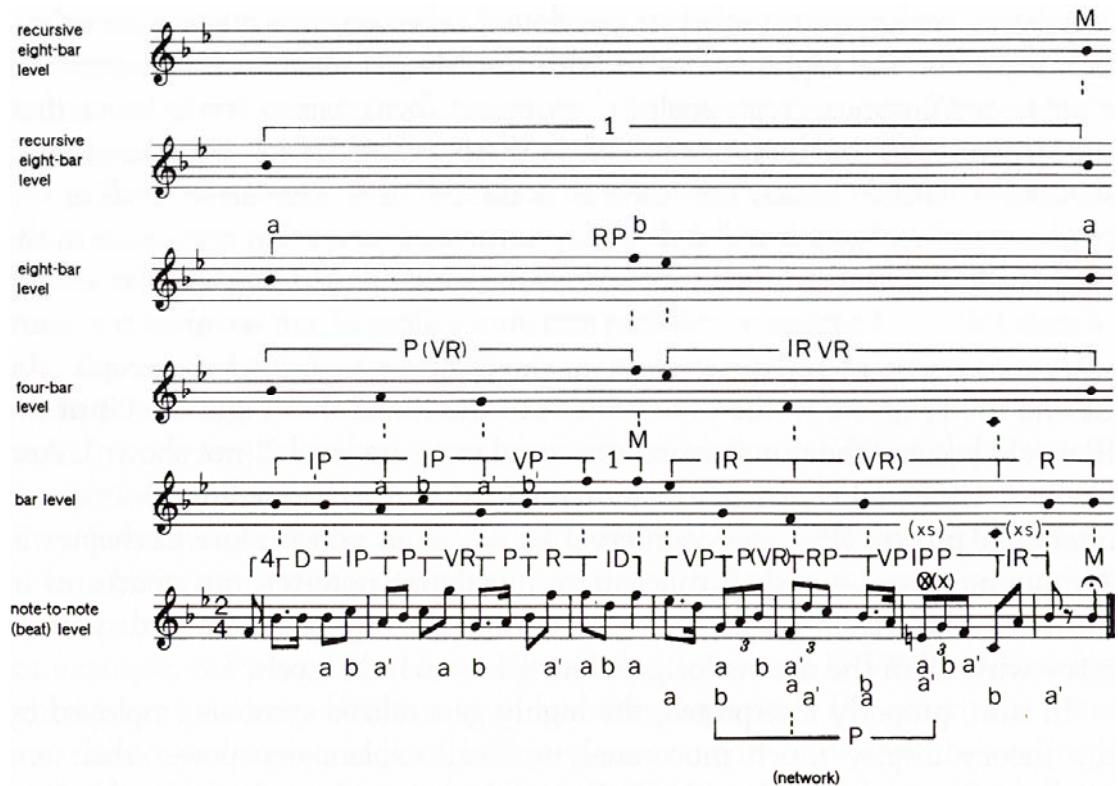
Figure I.6: Hierarchical analysis of a melody showing the basic structures of I-R model. It can be observed that the recursive application of the rules generates a derivation of successively higher-level structural tones extracted from the initial and the terminal tones on lower levels (see dashed vertical lines) (Extracted from Narmour 1992, p. 32).

I.2.3 Summary: prolongation as theorized

The three models reviewed here consider the problem of hierarchy in music to a different extent. However, what does it mean to say that music is hierarchical? What is a hierarchy exactly? According to Cohn and Dempster (1992) these questions face us with a problem of scope. In order to conceptualize music as hierarchical it is necessary to tackle the problem of the relative relevance of the features that inform a given hierarchy, and more precisely to consider the function of those features within the hierarchical system analyzed. We need to decide if, for example, a hierarchical structure is going to be understood just in

terms of the underlying compositional parameters that emerge from the unfolding of the voice-leading, or if the motivic characteristics of the musical surface will be taken into account as well. We also need to make a decision about which musical features have structural priority as factors to take into account in determining an underlying hierarchy; in other words, we need to decide whether it is the pitch component that holds primacy or there are other structural features (i.e. rhythmic) involved. Finally, it is crucial to consider the status of structural events within the underlying hierarchy. To be precise, are events literal representatives or are they metaphorical extensions of other events at different hierarchical levels? For example, do structurally important events of the musical surface become themselves underlying events that represent those *same* events at deeper levels of the musical structure (such as the notes A, B flat, C and A of the upper voice of Bach's choral in Figure I.4)? Or, different, are underlying events of a prolonging passage 'seen' as potential candidates that lead to some other states thanks to their capacity to compose-out linear connections (such as the prolonged linear descent formed by the notes G, F, E, D and C in the upper line of Haydn's piano sonata in Figure I. 2)?

Hierarchical structures imply the idea of predominance-subordination between events, where a relative importance is assigned to them according to contextual factors that are involved in the building-up process of the musical piece. As noted above (see section I.1, p.2), reductional models view underlying hierarchy as a unitary structure in which relationships between events are instantiated in terms of a one-to-many correspondence. Even though the music-analytic approaches that have been considered so far predicate, to some extent, this kind of one-to-many relationship between musical events, they approach it in rather different manners.

Prolongation, as an attribute of underlying musical structure, is one of the main constructs put forward in the field of music theory to account for the idea of a one-to-many relationship between the pitch events' organization of a tonal piece. The theoretical framework proposed by Schenker may very well be considered as the foundation of subsequent approaches to this concept. The principles developed in Schenker's theory about the underlying musical structure in general, and of prolongation in particular, are usually taken for granted in analytic papers, attesting to their strong contribution as explanatory tools of the architectural organization of pieces of music. Prolongation's value as an organizational principle that governs music's unfolding has been already recognized, but its worth as a cognitive principle for understanding the experience of tonal music is still in question (see Horton, 2003) and has scarcely investigated to date.

The treatment of prolongation in the theoretical sources reviewed here reveals ontologically different conceptions of underlying hierarchy in tonal music. The point is, how different is this "nominally identical" concept in each theoretical framework? Comparison of the ways prolongation is approached will highlight some issues relative to its ontology; clarification of these issues is necessary to investigate its cognitive reality.

The notion of *predominance-subordination* is invariably present in the analysis of prolongation. The relative importance of musical events in the organization that results from the application of this principle to the unfolding of the underlying hierarchy depends on their *degree of stability*. In Schenker's approach, stability is mainly restricted to the pitch domain. It is governed, at global level, by the coherence of tonal principles, instantiated in the archetypal organization of the fundamental structure, and at progressively local levels by the unfolding of voice-leading, via the elaborative processes prompted by prolongation. In *GTTM*, pitch stability is a necessary but not a sufficient

criterion to determine the degree of stability of the events' structure. Stability is governed, in the first place, by temporal constraints instantiated in the nested organization of the time-span segmentation. The time-span, therefore, becomes a regulatory factor that is used in the first place to determine the structural importance of an event; therefore, the time-span primes pitch structure considerations. In Meyer-Narmour's view, stability is governed mainly at the level of the musical surface by differences in the implicative nature of pitch-interval relationships.

Recursivity is another distinctive feature that characterizes the assessment of prolongation, and the theoretical frameworks reviewed here are enlightening about differences in the way the recursive procedure is applied to the analysis of the underlying hierarchy. *GTTM*'s recursivity is informed by a "science systems" approach. Hence, recursivity is the result of the operation of a system's organization in which every relation among events and/or collections of events is defined by two conditions: absolute disjointedness - each individual appears exactly once in the collection - and unique connection - whereby every element or node in a hierarchy is connected to every other element or node by exactly one path (Cohn and Dempster, 1992). Hierarchical sets recursively satisfy this requirement. Therefore, prolongational reductions in *GTTM* are constrained by a strictly nested, constituent hierarchy, and recursivity organizes the musical stream into syntactical, discrete portions of information, that is to say, into the constituent chunks generated by the time-span segmentation. By means of the application of the recursive procedure some distinct components of the segmented chunks are integrated into more comprehensive, though increasingly reduced, pieces of information. Important to our distinction between ontologically different types of prolongational hierarchies is that recursivity in *GTTM* proceeds by "selection" of *literal* pitch events at the level of the musical surface, in order

33

to generate, at the following higher level a new chunk which - in the nested hierarchy - is a *partial representative* of the musical content of the piece. This process re-applies until the whole nested structure is completed. In summary, *GTTM*'s approach represents the science system's assumption about the computational nature of the musical mind.

In contrast, Schenker's theory presents a transformational system in which the musical composition is built-up as an inclusional, embedded hierarchy of increasingly larger wholes. Structural pitch events that belong to a specific event-class, for example scale-steps, control musical spans all along the prolongational units, being understood as *abstract representatives* of the stream of events. Thus, prolonged events govern prolonging events but are not literally equivalent to them. In inclusional hierarchies the constituent events of a determined span are treated as equally distant to the prolonged event that controls the span, in order to preserve the absolute disjointedness of the hierarchy. Thus, for example, in a prolongational passage that embellishes a note, none of its constituent pitch events are closer to the prolonged event because, as prolonging events, they are all similarly contained within the passage. In this way, the problem of having to take arbitrary decisions about left or right branching, as is frequently the case in representational hierarchies, is avoided.

> "The inclusional ontology has the virtue of maintaining the hierarchical status of the model without imposing a series of arbitrary choices. At the same time, adoption of this view extracts a price: the relationship of prolongation to the traditional concepts of diminution and embellishments - and consequently to species counterpoint - must be viewed as metaphorical and heuristic rather than formal and literal" (Cohn & Dempster 1992, p. 163).

Therefore, instead of partially representing the musical content at each hierarchical level of the piece, an inclusional hierarchy preserves the musical organization of the composition through the transformations that take place in the underlying structure. By means of

prolongation, those events that are metaphorically abstracted control relatively longer musical spans, providing direction and coherence to the unfolding voice-leading. Given that the recursive procedure is not thoroughly followed in Schenkerian analysis as it is in *GTTM* and given the rather heuristic nature of Schenkerian analytical method it might be thought, as Cohn and Dempster in fact do, that there is a price to pay as a consequence. It is an open question, however, whether the heuristic nature of the Schenkerian procedure is an advantage or a disadvantage in accounting for the experience of certain aspects of the prolongational structure of tonal music.

Recursivity in Narmour's *I-R* is instantiated by means of closure, which functions as the mechanism whereby multiple levels are generated in the cognition of melody. Closure is defined in terms of the relative strength of melodic implications, and largely determines the way in which patterns of melodic notes may be grouped. The initial and terminal tones at lower levels are derived at successive higher levels, although according to Narmour "the extent to which listerners actually transform structural tones to such levels is still an open question" (Narmour 1992, p. 32).

There are also differences in the three approaches, concerning the ways in which the recursive procedure is applied in analytical practice. The *GTTM* and *I-R* models follow strongly constrained methods that guide the derivation of musical structures, including instructions about the directionality of the reductional process (from background to foreground and/or the other way round) according to generativity rules (*GTTM*) or implication-realization rules (*I-R*). In contrast - as noted above - Schenkerian analytical practice deploys a more heuristic procedure: Schenker's principles of the underlying organization of tonal music are continuously taken into consideration in order to obtain the voice-leading arrangement, but the derivation of the unfolding hierarchy takes into account

the notion that even though such principles apply throughout the underlying structure of a piece, each of the three hierarchical levels, that is to say, the foreground, the middleground and the background has, nevertheless, its own hierarchical substance (Schachter, 1999; see also Schachter's quotation above in section I.2.2, p.20).

Summarizing, *GTTM* tackles the reductional procedure in a different way from Schenker's theory. During the process of reduction between hierarchical levels, *GTTM* selects those tones considered structurally important, and, as a consequence of that selection, they become representatives of the tones at that specific reductional level. Underlying events in Schenkerian theory, however, *are not selected but metaphorically abstracted*; as a consequence of this process they control increasingly longer musical spans of the underlying unfolding voice-leading. In *GTTM,* hierarchies are "representational". In Schenkerian theory, they are "inclusional" (Cohn & Dempster, 1992). While in the case of the *I-R* model, although it also performs a kind of pitch-event selection, such selection is based on different constraints than the ones instantiated by the *GTTM*, being mainly restricted to the domain of melody.

Another critical issue pertaining to the ontological status of prolongation is the relative *priority* and/or *compatibility* (Cohn & Dempster 1992, p.165) assigned to those features that are included in the prolongational construction.

The principle of prolongational priority postulates that features other than pitch are dependent and determined by the pitch relationships of harmony and voice-leading, and not conversely; it follows from this that pitch features are prioritised. Schenkerian logic obeys this principle. Although contemporary Schenkerian theorists have elaborated on the prolongational dimensions of other aspects of musical structure, such as, for example, rhythm (Schachter, 1976) they still acknowledge the priority of pitch and its manifestation

in voice-leading. Thus, pitch controls the capacity of other musical dimensions to provide syntax in musical compositions.

The principle of prolongational compatibility posits that underlying hierarchies may be generated, constrained, or even controlled partly by features other than pitch, therefore accounting for the possibility of a mutual dependency. This principle applies best to *GTTM*. In this case other features interact with the voice-leading arrangement. The time-span reduction and the rhythmic priority in *GTTM*'s generative analysis, together with its capacity to prime the structures that result from the prolongational reduction procedure, is one example of the compatibility principle. The *I-R* model also focuses on pitch as the main source of structural priority although duration is taken into account as well.

As seen above (see section I.1, p.2), according to the view of contemporary music theory, musical unity and stability are two principles of tonal organization that are not exposed in the complexities of the musical surface but hidden in the underlying simplicity of musical works. In spite of the acknowledgement of the overall capacity of pitch to constrain the underlying structure of tonal music, so far there is little agreement as to the role that different dimensions of pitch organization play in conferring unity on the musical piece. Trying to overcome this apparent shortcoming, alternative explanations of musical unity have been recently considered (Cohn and Dempster, 1992). Rather than just constraining unity to the simplicity provided by the underlying hierarchy, these explanations reconsider its scope and understand unity as a by-product of the interaction between the constraints imposed by harmony and voice-leading on the one side, and the motivic-thematic relations that exist in a composition (Beach, 1984). Ambiguous surface motivic patterns may be clarified by means of their temporal expansion at middleground reductional levels (Cadwallader, 1988), and other types of motivic components that are not evident in the

musical surface (such as higher-level motives that supply melodic coherence in the unfolding voice-leading), and that are hidden in the underlying structure of the composition (Cadwallader & Gagné, 1992).

Some evidence of the potential interaction between alternative sources of musical unity can be found in the Schenkerian concept of motivic parallelism and concealed or hidden repetition (Burkhart, 1978; Rothgeb, 1983, Cadwallader, 1983). Through this idea it is posited that in a given tonal composition, a melodic unit or "motive" can be expressed (can be explicitly exposed) on different structural levels. According to Cadwallader and Pastille (1992) 'Schenker sees the presence of higher-level motives as a liberating force for the surface motivic relations' (p. 123). Buckhart calls such compositional features *motivic parallelism* (Burkhart 1978, p. 146). Sometimes, Schenkerian's analytical practice seems to support the idea of competitive interaction between different sources of musical unity, because, according to Cohn & Dempster (1992) in order to perform the recognition of such concealed repetitions through the underlying shape of the piece, the analyst is faced with the dilemma of being influenced by his or her *a priori* ideas about what and where to find those motivic parallelisms, or to perform the 'reading' procedure free of previous beliefs, following the rules of the underlying structure as the unique guide to solve the problem. In sum, this alternative ontology does not confine generativity to a single source (the harmonic or the melodic domain); it opens the possibility of admitting the interaction of a prolongational hierarchy with motivic or thematic relations, avoiding the problem of privileging either of them. Thus, "the identity of a neighbor figure depends not only on the existence of a particular sequence of pitches or intervals, but also on the structural inferiority of the interior event to the exterior ones." (Cohn & Dempster 1992, p. 168).

In spite of the number of music-theoretic sources that have questioned the concept of unity as a valid assumption for music analysis (see a recent debate on this issue in Korsyn, 2004; Agawu, 2004 and Kramer, 2004), it has been demonstrated that there are unifying components in tonal music. They can be observed, for example, during recapitulatory restatements of the theme, where voice-leading connections between the current musical passage and the formers are established, just before arriving to the closing cadence; or when ascending gestures in parallel motion lead to climactic points, in which the tonal cadence is articulated (just to mention two cases). Unifying components are linked to perceptible features of musical compositions; by means of their 'ultimate connectedness', they contribute to our understanding of 'what is remarkable, unique or surprising in a work' (Morgan, 2003, p. 42).

Schenker's theory of underlying musical structure has been and is still one of the main musicological referential frames in the analysis of tonal music. Its study and pedagogical applications (Komar, 1988; Beach, 1984) encouraged the development of a rich literature in the field of music theory, notably in the European-American academic milieu. During the second half of the twentieth century, Lerdahl and Jackenfoff's *GTTM* and Narmour's *I-R* models were created under the strong influence of the cognitive revolution, being deeply rooted in cognitive psychology and linguistics. Even though to some extent Schenker's fundamental postulates, among them the concept of prolongation, served as a basis to elaborate further theoretical formulations, as in *GTTM*, the new theories claim to have produced developments at the cognitive level that, to some extent, are intended to supplant 'traditional' Schenkerian principles. As the status of prolongation as experienced is examined in this thesis, it is necessary to consider in the first place the extent to which Schenkerian theory makes specific assumptions about its psychological nature and in the

second place whether the 'supplanting' theories - of Lerdahl and Jackendoff and of Narmour - are, or are not, adequate to the task of providing accounts of musical structure in cognition.

Schenker conceived his theory of musical structure mainly in response to his own aesthetic motivations. His theoretical conception has been subjected to extensive criticism for more than fifty years. During this time, and probably as a consequence of such scrutiny, it seems as if some of the criticisms that the theory faced and still does, do not appear to pay enough attention to the context in which Schenker's developments were formulated, dismissing essential aspects of his *idea* of music that are deeply rooted in the cultural environment of his time. In other words, what seems to be dismissed is the *conceptual model* (Zbikowsky, 2002) that is implicit in Schenker's thought (see Chapter VI for additional consideration of this issue) or, in terms of Foucault, the *episteme* that was dominant in Schenker's music-theoretical practice (Cook, 2002), including historic and ideological aspects that guided his thought (for an interpretation of Schenker's idea of "representations" as music-syntactical narratives, born out from the traditional genres of drama and epic, see Littlefield and Neumeyer, 1992).

Schenker's view of musical structure reflects in part the influence of aesthetic principles derived from the Classic philosophical tradition, refocused later by Kantian philosophy. According to Kant, the experience of form is the primary goal of aesthetic perception. Form is understood as an aprioristic configuration, of whereby image or sensation is elicited by the concrete work of art. This idea of totality is found in Schenker's notion of an underlying archetypal unity. The psychological component of Schenker's conception of musical structure is the result of the distinction he makes between art and nature: he differentiates between the physical properties of tones, acoustically measured in isolation,

and the formal properties of sound experienced in the context of the musical piece. Thus, what appears as acoustically equal in the surface is experienced as different in the deepness of the underlying musical context (Cook, 1987; Schachter, 1990). According to Schenker, the structural unity of a tonal piece implies more than a concatenation of tonal events, analysed comparatively in terms of relationships that, as *GTTM* posits, are intended to capture relationships of tension and relaxation, out of which some events are selected to be transferred to the next higher level, and so forth until the *Basic Form* is reached (Lerdahl, 2001). On the contrary, in the transformational system proposed by Schenker, an *a priori imaginative* representation of the underlying unity organizes the perception of the tonal structure, allowing the moment-to-moment recognition of the intelligible totality of the musical composition.

However, the point is whether these transformational principles model musical experience. It is an open question whether Schenker's idea of a large-scale structure awareness (that he accords, by the way, only to the genius of the great composers) is applicable to normal hearing. In fact, it has been severely criticized (see Narmour, 1977). As will be explained below in this thesis (see in Chapters III and IV the characterization of the basic level) it seems as if the most *natural* psychological level of apprehension of essential features of objects is a level that is neither very high nor very low, that is in the middle of the hierarchy. Therefore, it might be that Schenkerian *Fernhören* (meaning a kind of global structural hearing) could very well be experienced at intermediate levels of the underlying structure of a musical piece.

The point is to decide what it is that one hears in the music. According to Cook (1989) this is problematic, because the way an analyst listens to music is not necessarily similar to the way people normally hear music. The purpose of a musical analysis is not to reflect exactly

the way people listen to music, but to provide an explanation of the musical piece, the validity of which will depend on the extent it is accepted as a fine interpretation of the music's essence. Similar considerations apply to Schenkerian music-analytical representations (Agawu 1989), although their capacity to convey a symbolic picture of the grammatical element of a piece has been subject to criticism (Horton, 1999).

The practice of Schenkerian analysis in academia has emphasized its value as a tool that develops a *way of hearing* a musical piece that is acquired through training. Connected to this idea, the literature of Schenkerian analysis contains numerous indicators that convey the idea of prolongation as experienced. For example, there are statements that refer to cognitive assumptions such as "the ear interprets" (Salzer & Schachter, 1969; p.119); "the musical ear perceives" (Salzer & Schachter, 1969; p.123); "the ear connects" (Salzer & Schachter, 1969; p.135); etc. However, as was noted earlier in this chapter, these assumptions have been barely studied in terms of the cognitive processes the listener, either musician or non-musician, might employ in relation to prolongation (see, for example, Serafine, 1988; Serafine, Glassman & Overbeeke, 1989 for seminal experimental studies on the Schenkerian concept of underlying structure).

So, the question remains: is prolongation a mere theoretical tool, helpful in analysis, or is it also a construct that might be attributed a psychological status in music cognition? Recently, some efforts had been made to analyze the conceptual and/or experiential role of prolongation in music cognition. For example, Kielian-Gilbert (2003) presents a view of the experience of prolongation that combines issues of temporal processing with what she calls the degree of relative determinacy of theoretical (interpretative) vs experiential (perceptual) aspects of music apprehension. The idea of prolongation as experienced is described by Kielian-Gilbert as follows:

"Just as prolongation ('composing-out') describes a 'way of listening' that characterises the degree of willingness to hold onto one event while hearing another, it also includes the sense of the relationship of a later to an earlier event that enlivens that earlier event (and vice versa). It is a relational process that projects and embeds structural functions and part-whole relationships. To experience prolongation is to differentiate content at different levels of specificity". (Kielian-Gilbert, 2003, pp. 56-57).

She makes a distinction between prolongational and translational relationships in tonal hearing and analysis. The description of these relationships is approached using conceptual metaphors (see Chapter VI for a thorough explanation of conceptual metaphors): the prolongational dimension implies the experience of back-to-front-depth spatial listening orientation, while the translational dimension refers to left-to-right horizontal orientation of the linear succession of the musical piece. This analytical description is in line with new developments in the field of music cognition that point to the imaginative nature of conceptual understanding. Making assumptions about the ways in which underlying hierarchies might be apprehended cognitively is to understand the experience of such structures as the result of performing some kind of mental operations through which the listener builds a representation of the underlying structure of tonal pieces.

In this thesis it is assumed that the status of the prolongational structure as experienced is worth investigating and that the insights proposed by prolongational theory are open to experimentation. From the point of view of classic cognitive science there are aspects of prolongation as experienced that have not been explored to date. Although theoretical efforts have been devoted to constrain the idea of the underlying representation mainly to the domain of graphic analytical descriptions, there are nevertheless promising indicators of the manifestation of prolongation in music perception, mainly by the use of prolongational structures in some cognitive tasks (see Chapter III for a review of previous

findings). The series of experiments that will be pursued in the first experimental part of this thesis will continue on this line (see Chapter V below). On the other hand, later developments in the field of cognitive science that have made extraordinary progress in understanding the architecture of the mind have not yet properly addressed the question of the cognitive status of prolongational structures, at least in the terms proposed by Schenkerian theory. It is an assumption that some of the postulates arising from these developments are still worth investigating in order better to contextualize its cognitive scope. For example, Kantian idealism is more akin to alternative cognitive views of metaphorical thinking and imaginative listening than to those proposed within classic cognitive science. Consequently, feasible hypotheses relative to the manifestation of prolongation as imaginative thinking in music cognition will be pursued here: the extent to which prolongation is abstracted and/or imagined, and/or mentally retained will be explored with the aim of elucidating its status. The experimental work described in the second part of the thesis addresses the idea of prolongation as structural metaphor.

CHAPTER II

PSYCHOLOGICAL BASES OF MUSIC HIERARCHY

II.1. Psychological reality and evidence

During the last part of the twentieth century, the investigation of mental processes and mental representations (their nature, organization and functional roles) was a fundamental concern in cognitive psychology and, concurrently, models of hierarchical structures were created to explain empirical findings or to predict them.

The pioneering challenge to behaviourism came from Lashley, who presented a view of the problem of the temporal structure of human perception and action as extraordinarily complex and, so, as impossible to be reduced to a succession of stimulus-response units. In his seminal paper on the problem of serial order (Lashley, 1951), he suggested that speech and other forms of coordinated action must share common organizational principles, instantiating the existence of hierarchical processes in cognition. Lashley invoked hierarchical processing to explain that an adequate representation of complex series integrates sets of behaviours such as the motions of fingers on a musical instrument that occur too rapidly to allow for triggering of one act by its predecessor (Lashley, 1951). He also posited that mental organization should be understood as a dynamic and permanently active system that was a combination of many interactive systems (Lashley, 1951, p.135).

Theoretical models that interpret mental representations as being hierarchically organized present a picture of hierarchy in terms of a multilevel organization in which information is represented at higher, superordinate levels in a more general or simpler form, and, conversely, at lower, subordinate levels, in a more detailed manner. In the case of

hierarchical models of mental processes, similarly, superordinate levels control the operation of processes at subordinate levels.

Music psychology, in turn, was informed by such endeavours, and subsequently experimental research flourished around the idea of music hierarchy. When some psychological models describe the musical experience they also account for hierarchies in mental processes and mental representations. They assume that musical understanding is a covert, non-observable, mostly *unconscious process* that involves the listener's sensitivity to the variety of features of a musical piece. Through exposure to music, hierarchical mental representations - the products of the subject's estimation of the relative importance of those musical features - are developed. Should these mental processes and representations be investigated, their analysis cannot avoid issues of psychological reality.

As far as the problem of reality in multi-levelled music organizations is concerned, the status of the mental operations that take place when musical information is processed at different levels of the hierarchy is a key issue that needs to be subjected to careful scrutiny.

In the first place it is necessary to focus on the nature of the changes that occur when information *moves* between levels: this seems not only to be a question of the mere transfer of literal information between levels, but of *transformations* by means of which information located at a given level appears to be somehow modified at a different level of the hierarchy. Different types of information transformation may occur: abstractions, reductions, simplifications, among others. Transformations are different according to the types of hierarchies - processes or representations - that are being considered. For example, in multi-levelled organizations of processes, operations at a given level may promote changes in the nature of other processes, or control mechanisms at higher levels may be responsible for operations at lower levels (Broadbent, 1977, in Cohen, 2000). In multi-

levelled hierarchies of representations, on the other hand, representations at lower levels seem to inherit properties of representations at higher levels. For example, reductional music representations present a hierarchical organization in which higher structural events control lower structural events in such a way that the kind of one-to-many relationship that results from such an arrangement is recursively identified at different levels of the hierarchy (see below in section II.2 for descriptions of cognitive principles of music hierarchy).

Different hypotheses have been proposed to explain the transformations that take place in the representation of underlying musical events: for example, a structurally important event is selected from a musical passage and promoted to a higher level as a literal copy of the lower level event that it represents: this type of transformation occurs in the *GTTM* model. In other cases, the underlying event is mentally abstracted and considered to be manifest somehow at a more remote level throughout a given musical passage; it is understood as an abstract representative of a group of lower level events: this type of transformation occurs in the Schenkerian model (see Chapter I, p. 4). Concerning the way musical information representation *moves* from the surface to the background or the other way round, it was underlined in chapter I (where some of the hierarchical models developed by music theory were discussed) that it is still a debatable issue whether the flow of information and the control mechanisms that command it traverse the hierarchy in a one-way manner, or if there is a two-way interaction between levels (see Chapter I, p.15).

When the epistemological status of a hierarchical model is examined, several questions arise: do hierarchical organizations emerge simply as cultural constructs or conventions adopted by academics and researchers, or do they manifest themselves psychologically in terms of processes and representations? Furthermore, do they have neurological reality,

that is to say, a real existence in terms of brain organization? According to Cohen (2000) the difference has not been precisely established.

The history of science provides examples of metaphors that were created as imaginative constructs, which were not intended to draw exact psychological and/or neurological parallels, but which were nevertheless shown later to have cognitive reality as experimental, ontogenetic and neuropsychological evidence was found. Also, some cognitive models that originated as hypothetical constructs later became accepted, when evidence of their cognitive reality was found.

A central issue in assessing the status and reality of such constructs is the difficulty of undertaking experimental work related to non-observable entities such as covert processes and mental representations. Consideration of the existential status of these entities gives rise to an epistemological problem within the philosophy of science, due to the unavailability and/or isolation of models of cognitive organization which are required to facilitate their verification and/or falsification.

Cohen (2000) identifies four possible criteria for the assessment of the cognitive status of hierarchical models: i) *behavioural evidence*, provided by experimental findings in which results show differences in response times, error rates and quality of responses when hierarchical levels are brought under experimental control; ii) *neuropsychological evidence*, coming from case studies of patterns of impairment that show dissociations between different hierarchical levels and which also provide information about differences in the vulnerability to neurological damage; iii) *ontogenetic evidence*, consisting of observations of populations of children, in which it is possible to realise that some levels of the hierarchy are acquired before others, and that some levels are more resistant than others to the effects of ageing; and finally iv) *logical evidence* consisting of the development of

arguments by which a determined hierarchical organization proves to be advantageous and efficient.

A hierarchical model gains credibility and status if the predictions that it asserts have been rigorously tested: in other words, if it can plausibly explain experimental evidence. Credibility is obtained when experimental evidence cannot be plausibly explained in another way; that is to say, if evidence observes the principle of parsimony. This would imply efficient explanation. If those conditions are fulfilled, then behavioural, neuropsychological and ontogenetic evidence will provide results that can be explained by calling upon the corresponding hierarchies. But, as is usually the case in social sciences, in human sciences, and, of course, in art, evidence coming from a given phenomenon sometimes elicits alternative interpretations. In fact, the possibility of obtaining multiple responses during the experience of a given musical event is in the core of music's nature as a creative human activity.

The problem faced by research that focuses on the analysis of internal processes and representations is that there is no direct evidence of their existence, but there *is* indirect evidence to claim that hierarchical organizations have both psychological and neurological reality, thanks to *converging results* coming from different sources of experimental work.

In order to be acceptable, a hierarchical system must satisfy at least one of the four criteria of reality described above. For example, behavioural evidence is important, given that the psychological reality of a cognitive hierarchy could be seriously questioned were it not reflected in the performance of certain kinds of tasks. However, this evidence is necessary but not sufficient. Therefore, although none of the four criteria mentioned above is entirely conclusive in itself, there is a strong case when different types of evidence converge.

Of the four types of explanations proposed, logical evidence seems to be the most compelling, because hierarchical structures seem to be logically necessary for the performance of certain cognitive operations. In hierarchical models of processes, for example, control systems run the operation of mechanisms for planning, monitoring and decision-taking, involving higher-level processes that control lower-level actions. In hierarchical models of representations, hierarchical structures offer the means of accessing and storing information economically, and representations of factual knowledge at different levels of generality facilitate the identification of useful analogies between different types of processes and representations.

II.1.1.-The nature of cognitive hierarchies: characteristic features and psychological evidence

Cohen poses several questions about the nature of cognitive hierarchies that are interesting as far as the status of the cognitive reality of prolongational aspects of the underlying musical structure is concerned:

> Do hierarchical representations in different cognitive domains exhibit common features that confer special advantages? Are hierarchies fixed structures or are they assembled and modified as required? Is there an optimum level of specificity that is most appropriate or most efficient for storing, accessing, or communicating information? Does this level stay the same or does it shift with task demands or other factors? Does the optimum level of representation vary between individuals with factors such as age and education and neurological status? Is information at some levels more vulnerable to loss with the passing of time or as a result of dysfunction? When hierarchical representations are initially constructed does construction proceed in a top-down or bottom-up direction? (Cohen, 2000, p.5)

The answers to some of these questions will clarify aspects of music's hierarchical nature, and will set further avenues of research, in order to investigate the prolongational structure

of tonal music. In pursuing this search, concurrent evidence about its psychological reality will be obtained.

Should prolongation be understood as a hierarchical process, cognitive operations might take place at various levels, with higher operational levels controlling and modifying operations at lower levels. Alternatively, operations of prolongational processes might best be identified at an optimum, most salient, primary level. Concerning hierarchical representations, if prolongation exhibits aspects of self-containment at each hierarchical level (Schenker [1935] - 1979; Schachter, 1999), representations would involve the analysis of events hierarchies at single hierarchical *strata.* On the other hand, if prolongational organizations showed a common pattern that extended recursively all along the whole architecture of the underlying hierarchy (Lerdahl & Jackendoff, 1983) it would be necessary to develop a cognitive account of events' representations throughout the whole hierarchy.

Given the complexity of this problem, it is of course an extremely ambitious enterprise to try to find answers to all of these questions, and to provide different types of concurrent evidence coming from different sources, in a single piece of research. This goal could only be reached comprehensively through the development of a long-term programme of research. Consequently, the aim of the work described in this thesis is to provide a *partial reading* of the problem under investigation, tackling some of the issues relative to its status of cognitive reality, and providing some experimental evidence that might inform a more comprehensive theory of the prolongational structure of tonal music.

II.1.1.1 Approaching hierarchies: analytical dimensions

Several dimensions of analysis can be identified when hierarchies are investigated. First, multilevel organizations of cognitive processing can be analyzed in terms of the scope of the control mechanisms involved at different levels. Processes at higher levels seem to operate over longer periods of time, involving the whole organism and including goals and intentions. On the other hand, processes at lower levels are concerned with component actions and/or specific event representations. For example, there is some neurological evidence for the idea of executive control systems of behaviour, modulated by strong top-down activation. There is evidence of the neurological reality of the Supervisory Attentional System (SAS) (Shallice & Burgess 1996, in Cohen, 2000). Two types of control processing seem to take place during attention to music; they are related to the characteristic unfolding of musical events and involve the filling in of time intervals. It seems that the way music information is structured elicits two main attentional modes: one is more analytical and is related to the event-to-event sequential organization of the musical surfaces and the other involves expectations about the future course of events while the musical structure evolves in time (Jones & Boltz, 1989).

Second, the way in which the mechanisms controlling action sequences operate may illuminate aspects of their cognitive representations. In hierarchical models of action representation, these may occur before and after an action is performed, being useful in planning and guiding particular sequences of actions. Vallacher and Wegner's action identification theory postulates that action hierarchies exist as internal cognitive representations that may be externalized in overt performances of actions, serving to map one onto the other, and to provide mechanisms for planning, monitoring, and evaluating behaviour. They identify three groups of factors that potentially affect the level of action

representation: action difficulty, action context, and the person's experience with the action (AIT 1987, in Cohen 2000). Representations may be expressed at lower levels in terms of detailed physical movements, or at higher levels by means of overall planning involving goals or intended results. Relationships between levels are organized in causal and temporal sequences, with causes operating from top to bottom and consequences from bottom to top.

It appears that a given action has a dominant identity that defines the level at which the action is represented and monitored. Actions at higher levels, together with their corresponding identifications, provide the conditions for stability, because they allow goals to be maintained over time and with changing circumstances. On the other hand, lower level actions' identifications involve fragmented components that cannot necessarily be integrated smoothly.

An effective structural action-representation requires both the selection of optimal levels of action-identification, and the ability to shift dynamically between levels. Factors such as familiarity, complexity, experience and level of skill are assumed to influence the level of action-representation. Furthermore, social and situational contexts in which an event occurs may also affect the dominant level of representation.

Hierarchical action-representations give rise to a compelling argument that may be useful in the study of music representation as embodied knowledge. In the case of underlying musical structure, understanding musical events as *performing actions* at different levels of the hierarchy should provide a coherent, organic representation in which tonal goals are mentally maintained whilst there are changes in the contingent circumstances of music's development. In this context, prolongation, as a feature that describes the unfolding of the

piece's structure, may be hypothesized as *organic movement* that is governed by a given

course of action (Schenker [1935] - 1979; Larson, 1997; Lerdahl, 2001).

Third, the degree of stability of mental representations, according to the hierarchical level

that is involved, is another important aspect of the study of their psychological reality.

Overall, experimental inquiries of this concept show that results coming from hierarchical

processes and hierarchical representations seem to provide evidence of greater degrees of

stability at higher levels than at lower levels. However, according to Rosch *et al*. (1976),

there exists a *basic level* of hierarchical representation that is behaviourally and cognitively

optimal (see chapter 6 for a detailed explanation of the basic level). In spite of contextual

and/or environmental differences, objects at the basic level tend to be categorized as more

typical and categories are, as a consequence, more stable. In this optimal level of

representation, similarities between objects and/or events are clearly apparent and

differences are stripped away. Therefore, measures of similarity judgments are usually

employed when the structural typicality of objects is investigated. Similarity judgement

measures have been frequently used in musical experiments to study the perception of

melodic surfaces (Selfridge-Field, 1998). However, the structural properties that emerge

from the underlying musical organization would *also* elicit the subject's activation of an

optimal level of representation that is used in similarity judgement tasks. Therefore, the

application of similarity measures appears as a useful tool to obtain psychological evidence

of musical structure at the basic level (Medin & Ross, 1997).

Fourth, hierarchical representations of constituent structures imply the building up of a

mental organization in which boundaries between units become cognitive markers whose

status as cognitive reference points can be subjected to experimental investigation. It is

expected that information processing varies according to the structural importance of

information at different locations of a given constituent hierarchy. Interactions between bottom-up and top-down processing may also orient the use of attentional resources when different levels of a given constituent hierarchy are being processed. Reaction time responses appear to be a good behavioural index of the hierarchical status of segments in constituent representations. These measurements have been applied to provide behavioural evidence of the cognitive representations of constituent organizations in linguistic and music studies (Sloboda, [1985] - 1996; Stoffer, 1985).

Fifth, acquisition of different hierarchical levels in a given domain appears to be dependent on developmental concerns. In spite of the ontogenetic evidence accounting for age effects in the acquisition of hierarchies, controversy exists about the order in which they are acquired. For example, does children's conceptual development follow an upwards-downwards direction, from undifferentiated high-level categories to more specific low-level ones, as long as perceptual discrimination is present? Or does it proceed in the opposite direction, by gradual abstraction of similarities from specific categories to more general superordinate concepts? Alternatively, does representation at deeper levels of a given hierarchical organization follow a developmental trend in the sense that it is acquired at a determined age? Developmental aspects of the cognitive representation of musical structure have been investigated. Some studies provided ontogenetic evidence of the acquisition of structural hearing at about seven years old (Serafine, 1988). Other studies investigated the developmental trend involved in the cognitive representation of tonal pitch (Lamont and Cross, 1994). These results illuminate the interaction between cognitive mechanisms and environmental issues (enculturation/learning) in the development of the cognitive representation of structural aspects of music.

Sixth, the ways in which events are mentally organized into information packages, in the form of schemas and scripts, bring important evidence in support of the logical advantages of hierarchical structuring (Schank, 1982, in Cohen, 2000). Stored scripts and schemas are essentially generic 'packages' of high-level information that allow reconstruction of typical features of a given object or fragment of reality. Schemas are constructs that refer to one's knowledge about the world. They are general and structured entities used in comprehension, consisting of a frame that includes slots for particular information. Understanding, then, consists of filling in these slots. Scripts, on the other hand, are knowledge structures containing the sequence of events that take place in stereotyped situations. They are like schemas for routine events. Given their type of structure they are appropriate to make inferences while information is processed (Medin and Ross, 1997). The psychological reality of these information packages gains further support from the experimental evidence of its role in facilitating comprehension and memory in studies of information processing in different domains. Mechanisms of information-package activation have predictive power, especially for implicit comprehension. They are mentally stored in a way that is both economical and retrievable, as long as appropriate cues are activated.

The organization of events in musical compositions may also be understood implicitly in terms of simplified information packages in the terms proposed by Rumelhart and Ortony (1977). As far as musical structure is concerned, music psychology has developed a body of theoretical and experimental work around the idea of schemas as constructs that govern music representation (see for example Dowling & Harwood, 1986; Lerdahl, 2001; Lerdahl, 1991). However, there is another group of schematic structures that organise experience and understanding at the level of bodily perception, movement and imagination, which

Neisser (1976) and Johnson (1987) called embodied schemas; they are assumed to play an important role in structuring our experience. They are not fixed, empty, *a priori* templates for conceptualizing knowledge, but rather flexible plans that guide our expectations and anticipations in our interactions with the environment (Johnson, 1987). As previously stated, a musical piece can be experienced as a general course of events which is seen in certain dynamic organizations that emerge from different *musical schemas and/or scripts* representing *musical actions*. This way of interpreting music's temporal unfolding is assumed to be the result of the operation of the embodied schematic structures mentioned above, that are used to structure knowledge between different domains of experience. It has been proposed that dynamic aspects of mental organization are involved in the perception of art (Arnheim, 1984). Imagination, traditionally understood as related mainly to the domain of creativity seems to play a central role in the construction of embodied aspects of knowledge. *Image-schemas* (Johnson, 1987) the type of specific mental structures of embodied knowledge about physical relationships that take place at the basic level of our environmental experience would be involved in mental representations of structural aspects of music (see in Chapter VI a detailed explanation of the concept of image-schemas).

Seventh, memory for hierarchical organizations might account for the way both general and particular details of different musical compositions are remembered. Memory is an essential component in music processing. In order to make sense of the sonic stream organization of a musical piece, memory for events appears also to proceed hierarchically. So, it is expected that evidence of memory processing of music hierarchies will also provide clues for the study of their cognitive reality. Those models that account for the organization of the representations of an event postulate that representations of events in

memory invoke general knowledge from standard memory packages (see for example the concept of tonal hierarchy, investigated by Krumhansl in section II.2.1 below) and that non-standard aspects of particular situations are stored as specific indices that function as retrievers in the recall of instances of those situations (see below in section II.2.1 the concept of event hierarchies, by Bharucha). Even more general higher - level representations provide information in the form of intentions, such as, for example, success in achieving goals.

These high-level structures allow us to recognize similarities and analogies between superficially quite different events. Thus, for example, in the case of memory for a particular sonata form, cognitive processing may activate general schemas/scripts for introductions, developments and/or repetitions, as well as for the whole structure of the sonata form. Non-standard aspects of particular events, stored as specific indices would, in turn, be useful in retrieving those particular events from memory. Thus, a given episode and/or event can be stored at several different levels of generality.

Memory is an essential point to take into consideration if the aim is to find evidence of the cognitive reality of prolongational aspects of the underlying musical structure. Tonal coherence, as a feature of the underlying hierarchy of tonal compositions, is the result of a kind of musical arrangement in which the stream of events makes sense according to the constraints imposed by the underlying voice leading. As we saw in chapter I, these constraints are the product of general tendencies of movement and direction in trying to achieve partial and final goals. As far as those tendencies of the tonal organization of events are recurrent, then, memory for the general and/or particular ways the tonal arrangement appears within the composition will be the product of the listener's acquaintance with such general and/or particular modes of coherent organization. The

elaboration processes involving prolongation become specific ways in which those basic archetypal patterns that are stored in long-term memory as a result of the above mentioned idiomatic familiarity can unfold. The innumerable instances of variation that the voice-leading arrangement adopts as a solution to the unfolding of the standard underlying tonal component become indices in memory representation in invoking such fundamental, archetypal knowledge.

Eighth, experimental work has confirmed that lower-level information about hierarchically structured knowledge is more likely to be forgotten over time than higher-level information. Studies of the hierarchical representation of knowledge suggest that preferentiality exists in the way information is processed, and that this quality of processing serves to preserve important aspects of a given knowledge domain. In other words, 'level effects' appear to exist in recall. For example, there is experimental evidence that events that are higher in a given hierarchy will be remembered better than elements at lower or terminal nodes. It has been found in experimental studies that distributions of errors are closely linked to hierarchical levels, most errors occurring when items at subordinate levels were involved, fewer errors at intermediate levels, and the fewest at higher levels. These results show that higher levels of a hierarchy are more accessible and that information can either be recalled directly or be reconstructed inferentially (Sloboda and Parker, 1985). Subordinate information, on the contrary, seems to be more difficult to infer. It was also demonstrated a level effect when memory for musical sequences was investigated; it was found that the accuracy of recall was dependent on the assimilation of the musical sequence to an specific tonal alphabet, that is to say, a chromatic scale, a diatonic scale, tonic, subdominant and dominant chords (Deutsch, 1980).

In summary, there are several commonalities in cognitive processing and cognitive representation that can be identified when the status of reality of hierarchies in different domains of experience is investigated. Hierarchies have a dynamic quality that on the one hand give them the capacity to fulfil particular goals, through the use of pre-stored elements, and, on the other hand, to be revised and reorganized to fit shifting circumstances. This dynamic quality of hierarchical processes and representations applies specially to the analysis of the musical experience in listening situations: attending to the temporal unfolding of the stream of events of a musical piece requires participants to develop prospective and retrospective cognitive elaborations. By means of these they are able to use pre-stored information that guides their inferential capacity to anticipate the fulfilment of partial goals. At the same time, they are able to retain the existence of previously heard information, with the aim of interpreting the continuous incoming of incidental, low-level events. Finally, they will use this information to re-assess, revise and reorganize the more important, structural data in order to anticipate the future course of events that will lead to the fulfilment of more important, coherent tonal goals.

Overall, high-level representations give access to representations at lower levels but, conversely, in some circumstances, very low-level specific features act as pointers that give access to details of a given situation. This is the case in music cognition when specific features of the musical surface trigger the abstraction of higher-level processes in categorization activities of musical understanding (Deliège, 1996). Or in music composition when a motivic superficial feature of a musical piece is used as a compositional resource that guides the arrangement of the whole musical work (see Zbikowsky, 2002, p.34). Or, as it was described in chapter I, when hidden repetitions of motives that have been presented in the foreground are used to structure the unfolding of

the underlying voice leading (see in Chapter I, section I.2.3 a description of hidden repetition). It is a strong assumption that the listener *uses* that information to build inferences about the composing-out of musical coherence. On the other hand, in spite of the overall agreement about the quality of higher-level structures as frameworks that facilitate location and accessibility of information, it is not clear so far which hierarchical level provides an optimal access point. Nevertheless, it will be demonstrated later in this thesis that representations at the basic level, that correspond approximately to the commonalities of the tonal unfolding that are found at the foreground/middleground reductional levels of the underlying structure, appear to function as important cognitive reference points to process the underlying hierarchy (see below experimental work reported in Chapters V and VII).

Furthermore, the dominant level for cognitive operations changes according to task demands and also to other factors such as experience, familiarity and level of expertise. Even though differences between musicians and non-musicians are not usually found when basic musical cognitive tasks are investigated (see Bigand, 1990; 1994) it is well known that the ability to change the focus of attention consciously between levels of cognitive processing and representation is well developed in experienced individuals. What deserves further investigation, however, is the extent to which the focus of attention changes unconsciously, according to the structural demands of music unfolding.

Finally, it is still a matter of debate as to whether the acquisition of hierarchies follows a bottom-up or a top-down direction (for a seminal investigation in the acquisition of hierarchical levels in music cognition see Serafine, 1988). Although alternative explanations of the general-specific hierarchical approach have been provided (for example McClelland and Rumelhart (1986) propose a single-layer distributed model for a

description of the algorithmic microstructure of cognition), the cognitive model of hierarchy appears to be the best functional description of the macrostructural psychological level of cognition (see Cohen, 2000, p. 29).

II.2. The psychological reality of hierarchical structure in music: principles and models of music hierarchy

During the last thirty years a fruitful avenue of investigation of the concept of hierarchy, as posited by music theory, and how it is cognitively experienced, has given rise to a rich and consistent body of experimental research under the umbrella of the cognitive-structuralist tradition in the field of psychology of music (see Sloboda, [1985]- 1996; Krumhansl, 1990; Dowling & Harwood, 1986; McAdams & Bigand, 1994; Aiello & Sloboda, 1994; Deutsch, 1999; Howell, West & Cross, 1985; Howell, Cross & West, 1991; Deliège & Sloboda, 1997, for an overview). The cognitive-structuralist tradition presents a view of pitch processing and pitch representation in terms of a hierarchical organization in which several psychological principles are involved.

When hierarchical pitch processing and pitch representation are investigated, several dimensions of analysis are usually taken into consideration: First, if hierarchical pitch organizations are derived from the stream of events of the musical surface of a piece, those cognitive principles that are assumed to be involved in a hierarchical description of music's experience should to some extent be able to demonstrate a capacity to describe *the ways* in which some of the surface events should be considered *mentally* more relevant than others. Second, if the degree of relative importance of events is the main factor that accounts for hierarchy, then there should exist in the pitch events' organization of the musical piece some structural properties that make some of them *appear to be experienced*

as more relevant than others in music listening. Finally, if music unfolds in time, such that its experience is successive and irreversible to a certain extent, there must exist some form of meaning-making in ongoing listening situations that allows hierarchy to take place in the mind. In so doing, certain ways of cognitive processing should operate to allow interaction between what is known from before, and what is still unknown; that is, between *old* knowledge and *new* knowledge in the treatment of hierarchical information.

As long as these dimensions are taken into account, the cognitive experience of hierarchy seems to be the result of the interaction of two kinds of cognitive processes. First, it involves processing temporal information, with operations taking place at the note-to-note level of the melodic surface, with a strong influence of bottom-up processing; second, it involves processing non-temporal information - that is to say, operations with musical entities in which events are separated in time and connected across different temporal-spans, strongly influenced by top-down processing. It is from the latter processing mode, according to some authors, that the abstraction of hierarchical levels may emerge (Serafine et al, 1989), even though interaction between both cognitive modalities is always in play (see in Chapter III a more detailed explanation of Serafine's view).

In explaining the experience of relationships of predominance-subordination between musical events, a number of different criteria are called into play that will allow cognition to *organize the differentiability of events* within the musical surface*.*

In the first place, the acoustical dimension of sound interacts with our auditory mechanisms to provide a psychoacoustical framework from which those auditory principles that help stream segregation emerge. These principles (Bregman, 1990) bring the most direct evidence for the study of the linear component of the underlying musical structure, accounting for its psychological plausibility. The principle of stream segregation,

for example, seems to be closely related to the practice of the voice-leading arrangement in Western tonal music (Huron, 1991). Stream segregation includes a variety of psychoacoustic phenomena such as physical continuity, pitch proximity, integration, synchrony of attacks and differentiation of timbre (see a study of these phenomena concerning the principles of voice-leading in Huron, 2001). These become evident in the actions of both performers and composers, who apply them while manipulating different acoustic parameters with the purpose of arranging structural attributes of the musical piece. The analysis of those practices reveals that they should not be understood just as mere conventions, derived from stylistic considerations, but as the necessary consequence of their subjection to principles of auditory organization that make use of universals of aural perception. Thus, for example, the methods and rules of Renaissance counterpoint, which at first might have had the aim of controlling the sensation of roughness associated with the dissonance, later acquired a syntactical role in the configuration of the musical style.

It is probable, then, that procedures ruled by strict counterpoint, from which the foundations to understand prolongation from the perspective of voice-leading unfolding are derived, turned themselves into a way of controlling for dissonance and fusion in order to preserve the independence of the voices (Huron, 2001). Accordingly, the tendency to elaborate the underlying voice-leading under the rule of stepwise-pitch linear movement may also originate in these psychoacoustic principles: "if the listener learns that the line is in a determined region, he learns to segregate it, no matter the presence of other elements in the texture" (Bregman, 1990, p. 516). Summarising, it is assumed that this mechanism, used to highlight structural attributes with the purpose of fulfilling communication goals, is also used by listeners to understand linear aspects of the underlying musical structure. The

insights to which they give rise become valid explanations of the working-out of such musical structures.

The varieties of psychoacoustical problems that emerge from these linear arrangements form the bases for further cognitive-structuralist considerations of hierarchical pitch processing. Thus, for example, issues of consonance-dissonance, contextual and/or inherent stability (Larson, 1997), segregation of the linear component, and psychoacoustical resistance of some pitch-events to be assimilated by others within the sonic stream of events (Benjamin, 1982) among others, contextualize the cognitive processing of prolongational aspects of music structure. A more comprehensive account of these factors is nevertheless beyond the scope of the present thesis.

Second, in tonal music it appears that both the frequency of occurrence of events that involve statistical counting of pitch-event durations, and also their organization in those recurrent patterns that characterize a particular musical style, are factors that facilitate hierarchical processes and representations. However, we will see later that the way in which these factors operate is tackled in a rather different manner, depending on whether the cognitive model approaches the cognition of music hierarchy according to a more static or a more dynamic perspective (see sections II.2.2 and II.2.3 for descriptions of static and dynamic models in music cognition).

Third, to the extent these different ways of processing interact, they allow the listener to contrast the stream of information of a given piece with the internalised idiomatic knowledge of the familiar musical culture is surely one of the main features that constrain the cognitive processing of music hierarchy.

Differences can be found in the ways that psychological approaches tackle the operation of these factors but, in general, all of them somehow recognize the existence of those cognitive dimensions. Reviewing the large body of research concerning the cognition of tonal music, it is possible to find evidence of several cognitive principles involved in musical structure processing, that may support the psychological plausibility of prolongational structure. What follows is an account of three of the main principles so far investigated.

II.2.1 Some cognitive principles of hierarchical structure in music

II.2.1.1 Tonal hierarchies and event hierarchies.

A basic idea that has guided research in music cognition, previously applied to other cognitive domains such as vision, for example (see Marr, 1982 for a review of highlights in vision investigation) is that the perceptual system assigns to certain members of a given category a special status, configuring them as cognitive reference points against which other category members are judged (Rosch, 1975; Rosch and Mervis, 1975). Application of this principle to the analysis of tonal pitch events results, by analogy, in a hierarchical description that will assign a central importance to one of them, the tonic, that is most typical, and specific positions to the rest of the tonal pitch events, according to their functions relative to the tonic. Thus, typicality is understood as the degree of stability. Most typical events are perceived as more stable and, conversely, less typical events are sensed as more unstable. The relative stability of different tonal pitch events is hypothesized in terms of cognitive proximity, and cognitive distances are represented in a geometrical space (Krumhansl, 1979; 1990; Lerdahl, 2001; for a detail explanation of geometrical space representations).

Familiarity with a multiplicity of musical compositions results in the abstraction of an invariant hierarchy of relative stability of pitch-class events. This type of implicit knowledge, called tonal hierarchy (Krumhansl, 1990) is acquired through repeated exposure to the regularities of a given musical style. Tonal hierarchy is a cognitive organization that contains information about the stability of every event member of a given pitch-class organization, and which therefore contributes to the perceptual stability of the pitches of a musical piece (see Figure II.1).
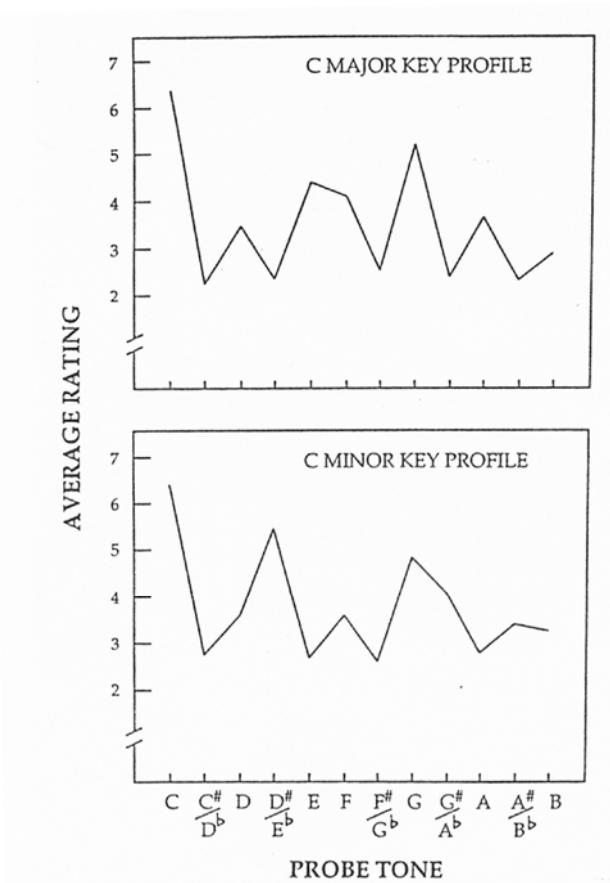
Figure II.1 Standardized probe-tone profiles for the C major and C minor keys. It can be observed that each pitch has a different average rating that corresponds to the perceived schematic tonal organization and that pitch events that are perceived as structurally more important are rated higher. Extracted from Krumhansl and Kessler 1982, p. 343.

Tonal hierarchies are stored in long-term memory in terms of schemas (see a description of schemas above in p.56). Once a given tonal hierarchy has been activated during listening to a composition, it contributes to the expectation of more stable events, helping encode every incoming event by contrasting its contextual stability to the inherent stability of the hierarchical pitch organization.

A tonal hierarchy is different from another hierarchical representation, called an event hierarchy (Bharucha, 1984b). The latter is formed by the specific pitch events of a certain musical composition (tones or chords) and contains information relative to the function of each pitch event's appearance in the piece. Each occurrence of the same tone, eg. F, in the context of a given composition represents a distinct presentation of this pitch event, in the sense that it can be experienced differently according to the context in which it is operating. Each musical piece has its own event hierarchy; however, most musical pieces of a given culture refer to a small set of tonal hierarchies.

Tonal hierarchies have been experimentally tested for pitch-class organization (Krumhansl, 1979; Krumhansl & Shepard, 1979; Krumhansl & Kessler, 1982) and for chord-class organization (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha & Castellano, 1982) in relation to an induced tonic (the whole experimental work was summarized in Krumhansl, 1983; 1990). These studies gave rise to a validated hierarchy of pitch-class events, called probe tone ratings. They were subjected to a multidimensional scaling procedure that produced a spatial pitch events representation which can validate theoretical descriptions produced by music theory concerning the organization of the tonal system. It appears that information about order effects within the string of events, and of the distribution of tone durations in the organization of a given event hierarchy, influence the activation of a specific tonal hierarchy (Bharucha 1984b). Tonal hierarchies provide a

kind of static, invariant, internalized knowledge of the pitch-class relative stability of the tonal system. They are assumed to be invoked each time the listener processes an event hierarchy, in order to compare the stability of each incoming event of a musical composition with the tonal invariant stability stored and represented in long-term memory; they also facilitate generation of expectations for the most stable events to occur.

In summary, tonal and event hierarchies are both psychological phenomena in which cognitive reality is supported by behavioural and ontogenetic evidence (Krumhansl, 1991; Lamont & Cross, 1994). They account for explanations of music representation in terms of the relative stability of pitch events and are also informative of hierarchy in terms of relations of dominance-subordination between events. However, the extent to which these two principles provide a convincing picture of an internal representation of the underlying unfolding of the musical structure of a tonal piece has been subjected to criticism (Horton, 2003). A tonal hierarchy is theoretically codified as, for example, the major and minor modes in Western tonal music (Dowling, 1984) and assumed to be mentally established as a hierarchical representation of the tonal functions that these modes elicit in the minds of competent listeners each time they are exposed to familiarity of the musical idiom. But it is arguable that those *are* the kind of mental representations that competent listeners *coherently* evoke when they listen to tonal music. One is inclined to think, for example, that the internalization of the underlying structure should imply more than just the implicit knowledge of a static picture conveyed by statistical counting of the distribution of event durations contrasted with a static invariant hierarchy of pitch-class tonal stability. That is to say, tonal and event hierarchies are psychological principles that provide a necessary but not sufficient explanation of the problem of how the underlying hierarchy of tonal music is mentally represented.

If a hierarchical interpretation assigns some events structural priority over others, and if such interpretation, in order to be realized, needs to rely on an invariant generalized abstracted pitch event organization, then it will always be the case that a given event will be assumed to be subordinated to another event that is more stable (for a more comprehensive explanation of the concept of structural priority see Chapter IV). But we know that this is not the case in tonal music, because what counts as important is not only how inherently stable is each scale-step, according to a ranking, but mainly what structural roles are played by scale-steps within a given context; in other words, what is important is how events relate to each other; the structural roles they fulfil in relation to one another. Therefore, stability cannot be the single criterion which regulates reduction as a cognitive process of hierarchical representation. In any case, it would be more appropriate to talk about contextual stability (Larson, 1997). Each appearance of a given pitch event does not, phenomenologically speaking, convey the same representative picture because, depending on its own contextual relationships, it is the concurrent interpretation that the mind is going to derive. An F is not always the same F.

II.2.1.2 Melodic anchoring

The principle of melodic anchoring characterises the listener's implicit tendency to assign to each incoming tone a position of relative stability or instability with regard to the tonal organization of the piece of music. It is a psychological principle developed by Bharucha (1984a; 1996) which operates on the musical surface of the piece, by which an unstable tone is assimilated to the tonal schema (Dowling & Harwood, 1986). In melodic anchoring, a dissonant event in a stream of sounds conveys a dynamic quality that induces an expectation of resolution towards the following consonant event. Once a tonal hierarchy (Krumhansl, 1990) has been established or activated from long-term memory,

the unstable tone interferes with the remaining tones until it is resolved. According to this psychological principle, the governing perception of an event's resolution is subjected to the operation of syntactic relations of order between pitch events that are responsible for the need of an unstable tone to be resolved by the following stable neighbour tone, in order to fulfil schemata assimilation. Two constraints govern the activation of this psychological principle: (i) the asymmetric property of the succession of events, in which an unstable tone must always be followed by a stable tone in order for the former to be assimilated by the latter and not the other way round, and (ii) the event's pitch proximity, meaning that tones must be proximal in pitch (Deutsch, 1982).

Although to some extent melodic anchoring is related to principles that manage the process of abstraction of structural aspects of a musical piece, the scope of its operation is strictly constrained to the note-to-note level of the musical surface. Thus, melodic anchoring operates rather differently to the principles of the underlying voice-leading. First, processes prompted by melodic anchoring are prospective, while reductional processes compromised by melodic diminutions (one of the voice-leading reductional techniques), for example, are either prospective or retrospective. In other words, we can find prefix and suffix diminutions (Forte & Gilbert, [1982]-1992; Lerdahl, 1997; 2001). Second, temporal and pitch constraints result in a view of linearity that is different in each case. While in melodic anchoring linearity is mainly the consequence of temporal order and pitch proximity between two successive tones, in the case of the voice-leading arrangement, application of linear principles derives in an organization in which less structural, surface events (prolonging tones) are related to other more structural events (prolonged tones) that are relatively distant, temporally speaking, from one another. Although in voice-leading organizations step-wise linearity between two structural events

is also assured, it generally occurs at more remote levels of the underlying structure.

Third, anchoring theory states that in two-tone relationships, an unstable tone (a non-chord tone) is resolved in a stable tone (a chord tone). Voice-leading theory, on the other hand, suggests that the second tone is not just a chord tone but a tone capable of establishing with the former a coherent horizontal relationship at a more remote level. For example, in the case of melodic diminutions, voice-leading principles assign different structural status to consonant skips and arpeggiations, while from the point of view of melodic anchoring they both represent cases of chord-tone relationships and therefore are assigned the same structural status.

Nevertheless, melodic anchoring is a valid principle that operates as a foundation of those psychological processes which are elicited by the implicative nature of the note-to-note relationships involved in melodic expectation (Meyer, 1973; Narmour, 1990; Larson, 2004). It sets up the basis for further recursive processing that takes place at higher levels of musical structure, although its value as a principle that proves useful to account for pitch processing at more remote levels than the note-to-note level of musical structure deserves further investigation.

These principles can be identified as cognitive mechanisms that inform more comprehensive descriptions provided by some of the main cognitive models of hierarchical processing of music. These models are reviewed in the section that follows.

II.2.2. Static models of hierarchical processing

The insights provided by the cognitive-structuralist literature on music cognition support the overall view of the psychological plausibility of the underlying music hierarchy. In this section we will pay attention to some of the main highlights that are specifically related to

the focus of the present study, other aspects that would be included in a more comprehensive account of the cognitive-structuralist view being out of the scope of this thesis.

Some psychological models base their assumptions on the 'systems science' tradition and the metaphor of the computational mind. They claim that experienced listeners process musical structure in a reductional mode. According to this, musical cognition is the result of a series of mental operations in which musical representation occurs by processing the stream of events of the piece in terms of previously acquired information of a musical idiom, abstracted and stored in long-term memory. Idiomatic musical knowledge consists of the stylistic regularities of that idiom, a kind of information that may or may not be explicitly exposed in the piece that is being heard, but that may be inferred by the listener.

Following former developments in the theoretical modelling of hierarchies of operators (see, for example, Simon & Summer, 1968), Deutsch & Feroe (1981) developed a logical formal model designed to account for the processing of pitch structural information in terms of the kind of reductional encoding that is involved when we listen to, memorise and reproduce tonal melodies. In this seminal model of music cognition, encoding of tonal sequences is represented by means of two components: hierarchically organized *alphabets*, typical modes of organisation of pitch events of the tonal language (octave key; triads and seventh chords, diatonic scale and chromatic scale, ordered from the superordinate to the subordinate levels respectively) and *formal operators*, which capture the type of relationships produced between elements of a given alphabet (ascent, descent, repetition, and distance, measured in number of semitones, between successive pitch components). According to the authors, formal operators account for the "internal design" of musical patterns, i.e. the internal composing-out of the elaborative melodic processes according to

73

which musical structure unfolds at a given level of the hierarchy, while alphabets account

for the structural frame that supports the voice-leading arrangement of the musical piece.

Principles of parsimony in melodic arrangement are related to the facility of the listener's

encoding. Alphabets, in Deutsch & Feroe's model, are organized in a reductional way.

That is to say, one level elaborates the content of its immediately superordinate level.

Experimental work on this model has provided evidence of its operation using melodic

sequences (see p. 94 in this chapter a description of the experiment).

Lerdahl's *Tonal Pitch Space* (1988, 2001), a further development of the ideas of *GTTM*

(see Chapter I for a description of Lerdahl and Jackendoff's *GTTM* (1983), posits a strong

psychological assumption about the way hierarchical musical structure is represented. The

model, which integrates music cognition and music analysis, works on the assumption of a

spatial representation of music: a metaphor that is not new in the fields of music theory and

music psychology (see Longuet-Higgins, 1987; Shepard, 1982).

Lerdahl's model synthesizes the contributions of a number of theoretical and experimental

sources of evidence in the domains of music cognition and music theory. It develops the

distinction drawn by Bharucha (1984b) between two psychological principles, tonal

hierarchy and event hierarchy (see in p. 66 above an explanation of these two concepts).

Tonal hierarchies are paralleled with *GTTM*'s stability conditions (see Chapter I, section

I.2.2.2) in the sense that they represent permanent, atemporal internalized knowledge that

goes beyond particular pieces, applying to entire systems. According to Lerdahl, event

hierarchies, on the other hand, deal with hierarchical relationships inferred from the

sequence of events of a given musical piece. Lerdahl considers them as a kind of pitch

reductions; in his model, prolongational structures are treated as forming event hierarchies.

We see that this is not exactly the meaning given originally by Bharucha to this concept,

nor does it mirror the way in which underlying voice-leading arrangements operate in Schenkerian contexts.

Tonal pitch space takes Deutsch and Feroe's (1981) algebraic spatial representation as the basis for the development of its own view of spatial representation. Deutsch and Feroe's hierarchically organized alphabets have two advantages for a pitch-space theory: first, although they are representatives of the tonal hierarchy, they take an algebraic rather than a geometric form, and second, the alphabets are organized like a musical reduction: the content of one level elaborates the content of its immediately superordinate level. *Tonal pitch space*, in turn, presents in the first place the *basic space*, a rendering of Deutsch & Feroe's hierarchically organized alphabet's representation, and goes further in its development, providing measures of pitch proximity. Quantification of distances between pitches operates successive transformations in the basic space, creating an algebraic representation that incorporates three tonal levels: pitch classes, chords and tonal-harmonic regions. The resultant representation is a stratified hierarchy in which the more stable elements at one level repeat at the next larger level. The model represents the relative proximity between all elements at each level and shows how the levels interconnect with each other (see figure II.2).

DIATONIC SPACE

a)   Vertical distance from pc0:   0   4   3   4   2   3   4   1   4   3   4   3

           Pitch class: *0   1   2   3   4   5   6   7   8   9   10   11*

b)   Horizontal distance   0   1   1   (0)

           from pc0:   0   1   2   3   3   2   1   (0)

                            0   1   2   3   4   5   6   5   4   3   2   1   (0)

           Pitch: *0   1   2   3   4   5   6   7   8   9   10   11(0)*

c)

d)   Combined distance: 0   5   6   4   6 6   3   5   7   6   2   6   7   5   7   6   4

           Pitch: C   D♭/C♯   D   E♭/D♯   E   F   G♭/F♯   G   A♭/G♯   A   B♭/A♯   B
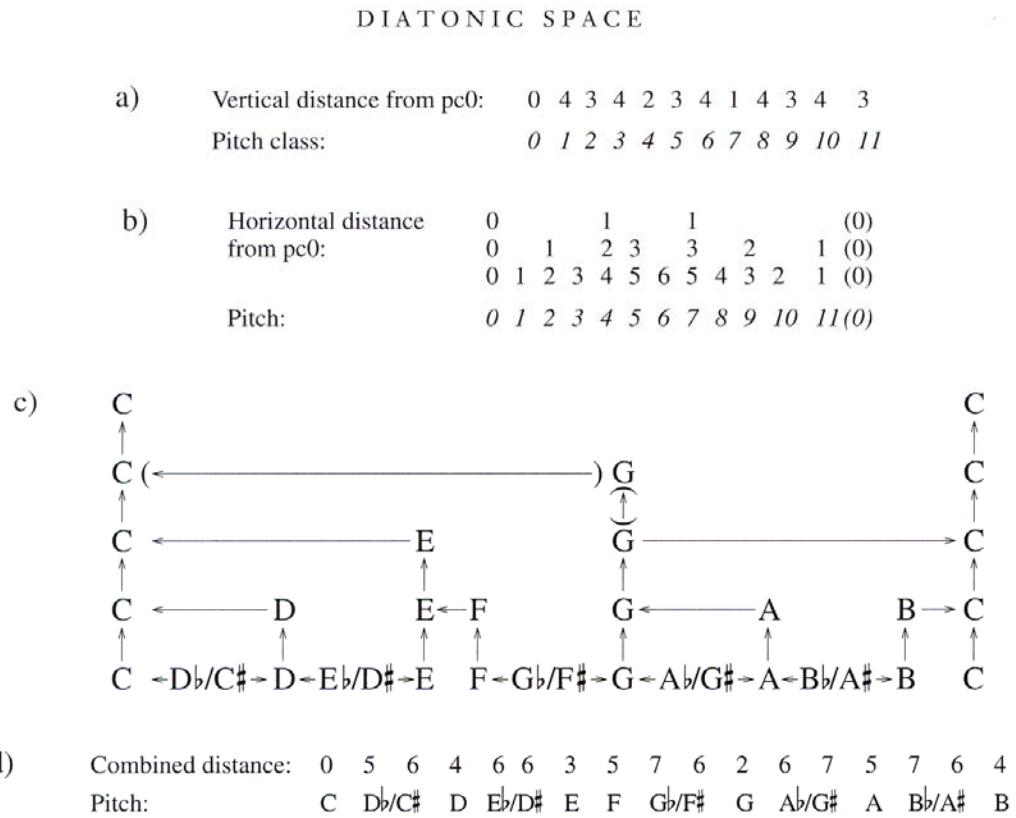
Figure II.2 Basic space representation of TPS. Pitch class and pitch proximity are representing in the following way: a)vertical depth of pc embedding; b) number of horizontal steps for each level; c) stepwise horizontal and verical pitch paths; d) number of moves for each pitch in c). Extracted from Lerdhal, 2001, p. 49.

Lerdahl's quantification of pitch proximity includes Krumhansl's experimental results on pitch, chord and region's proximity and her consequent geometrical solution (Krumhansl, 1990). *Tonal pitch space* calculates pitch proximity both in the vertical (from each pitch-class appearance in relation to the tonic) and horizontal dimensions (in terms of steps and skips). The interesting point here is that pitch proximity, in terms of steps and skips, is treated from a cognitive point of view, taking into account the hierarchical level that is considered in the analysis. Thus, for example, in the triadic level, C-E and E-G are steps, while C-G is a skip. That is to say, steps in each level are unequal in log frequency but cognitively equal.

Tonal pitch space explains the linear component according to the principles of linear completion and the shortest path. Linear completion is regulated by the constraints imposed by dissonance and its relation to deepness in the spatial distribution: moving down in the space brings more dissonance. The principle of shortest path handles operations of passing and neighboring pitches that are related to the anchoring principle from Bharucha (1984a) (see above in p. 70 an explanation of this concept). It states that resolution of dissonance between pitch events is due to their more consonant neighbours, in particular by subsequent stepwise resolution. Melodic closure is described in the model as a dissonance resolution on a larger scale; it is attained when the last member of the sequence is superordinate within the space.

Tonal pitch space also applies the implication of continuation principle (used by Meyer, 1973 and Narmour, 1990) to analyse the note-to-note pitch-step level) to regulate linear completion in other hierarchical levels; for example, continuation in the triadic level. Relative melodic closure, understood as the relative absence of implication, occurs when a more stable melodic point is reached (see chapter I for an explanation of Narmour's *I-R* model).

The analysis of linear completion features provided by Lerdahl's model can be paralleled, at a certain extent, by the voice-leading procedures of linear progressions (*Zug)* offered by Schenkerian theory. A linear progression is defined in Lerdahl's model as an upward or downward diatonic stepwise progression bounded by pitches of the prolonged harmony, operating at any level in an event hierarchy (see Chapter I for differences between this approach and that offered by Schenker's model).

In synthesis, linear completion in *tonal pitch space* consists of one-directional stepwise motion at any level in the basic space, such that the boundary pitches of the line are

superordinate in the space. The end points of a complete line can also belong to a level more than one stage above the level in question. Completion occurs not only at the musical surface of a piece but also at underlying reductional levels. Therefore, linear completion is viewed in *tonal pitch space* as a fundamental voice-leading principle of tonal music.

Finally, the principle of shortest path states that the pitch-space distance between two events is preferably the smallest value. It provides a crucial link between pitch space and prolongational structure. The preference for the smallest value means that the listener prefers the most stable choice, and the most stable choice is the one that is closest within the superordinate context. This principle replaces *GTTM* stability conditions and resembles intuitions provided by Schenkerian theory. Application of the principle of the shortest path originates the tracing of pitch paths all along the tonal space at multiple prolongational levels.

Computing pitch space distances in terms of their tension/relaxation unfolding of musical patterns, and modelling the way listeners operate with these variables according to pitch expectations, *tonal pitch space* intends to provide a formal, although rather static explanation for a tonal theory of music representation.

However, one of the major dilemmas within music psychology, in trying to find appropriate answers to the problem of musical experience, concerns the way listeners make sense of the ongoing music unfolding in real time. Some cognitive models tackle the problem of the sonic organization in terms of a continuous temporal input, and focus on the mental processes that are supposed to take place when the listener deals with the stream of events of a musical composition. In the following section we review two outstanding approaches.

II.2.3. Dynamic models of hierarchical processing

Jackendoff ([1992]-1999) developed the *Parallel Multiple Analysis (PMA)* model, a cognitive approach of musical processing inspired by the psycholinguistics tradition. According to his view, a piece of music is a mentally construed entity, scores and performances being only partial representations through which its meaningful reality is communicated. The theory models the connection established by the listener between the musical surface of a piece and the structure he attributes to it. "The perception of music (...) involves the unconscious construction of abstract musical structures, of which the events of the musical surface are only the audible part. The abstract musical structures so constructed are what account for one's musical understanding". (Jackendoff, [1992] - 1999, p. 51). According to Jackendoff, therefore, musical structures are unconsciously constructed as abstractions of the events of the musical surface. He understands musical representation as a by-product of the listener's internalized familiarity with a musical idiom. But he also hypothesizes that representation is built by listeners in real time situations during the process of *parsing* a composition (in psychological theories of speech processing, parsing is the process by means of which the listener assigns syntactic structures to sentences). That is to say, understanding hierarchical structures occurs in ongoing contexts, as long as the listening activity progresses in time (Jackendoff, 1987; [1992]-1999). *PMA* combines the operation of two cognitive processes: prospective listening, which generates expectations, and retrospective listening, which allows revisions of the expectations previously generated without compromising the speed of processing. In order to process the on-line structure of the musical piece, its musical surface is parsed and several analyses of its ongoing structure are pursued simultaneously. However, given that the parser can attend consciously to only one analysis at a time, the

other possible analyses are unconsciously performed, being inaccessible to attention. Jackendoff proposes a selection function in the mental processor that allows choosing the more plausible structure that the listener *is* aware of at a given point of the parsing process.

Given that the cognitive principles revised in section II.2.1 are some of the most basic mechanisms that operate at the note-to-note analysis of a musical stream of events, and given their efficacy in providing psychological priming to pitch processing at that musical level, they may well be some of the mechanisms involved in eliciting any of the various alternative-parsing hypotheses pursued by the listener.

It is a strong assumption in this thesis that underlying representations, as they were considered before (see in chapter I the concept of representational and inclusional hierarchies), are informed by the relationships between the internalized familiarity with a musical idiom and the prospective and retrospective quality of the cognitive processing that is unfolded during ongoing listening situations. In order to realize metaphorical extensions it is necessary that the mind proceeds consistently, in a forward-backward manner in order to turn the potential and implicit parser's analyses into reality (see Chapter I, section I.2.2.1 for a description of the concept of metaphorical extension). However, investigation of the ways in which mental processing of musical structure is deployed over time requires a solution that depends crucially on the interaction of attentional resources of the human mind with the listener's internalized knowledge of the musical idiom.

Jones (1992; 1993) and Jones & Boltz (1989) developed a model of the way attention is controlled by structural aspects of Western tonal music. This approach is relevant to our investigation of the capacity of prolongational features to prime the way in which

attentional resources are used to process the temporal unfolding of the underlying

hierarchy. The model hypothesizes that structural components of a musical piece may

"orient" the way in which the process of attending takes place (Jones and Boltz, 1989).

Temporal responses seem to be based on the analysis of the organisation of events, and the

dynamic aspects of attending. The model assumes that event structure defines time

intervals, affecting the way people tend to judge an object's temporal dimension.

Temporal extrapolations related to musical pitch organization can be derived from this

approach. According to them, the temporal unfolding of music should be influenced by a

variety of non-temporal factors, such as the distribution and complexity of musical events,

which interact with our attentional disposition. The way musical patterns unfold affects

estimations of the completion of temporal intervals, orienting our attention to the future

course of events. Thus, for example, two temporally equivalent intervals will be judged

differently depending on the non-temporal information contained in them (Jones and Boltz,

1989).

If information-processing is a function of the surface structure of the musical piece, there

should be a relationship between the underlying structure and the way attention is

deployed: the latter will vary according to the structural importance of the musical event

that is being focused upon, and decisions about what is actually heard will be related to the

unfolding of the underlying organization of the piece.

Jones' proposal tries to overcome a series of constraints identified in previous models of

attention, including (i) their exclusive consideration of non-temporal constraints, with the

concurrent neglect of the impact of temporal information as a referent to produce judgment

durations; (ii) the lack of clarity in providing a precise definition of psychological

complexity as a factor to determine, for example, the extent to which a chunk can be

experienced as such and what typical tasks determine chunks' boundaries (this issue will be treated in detail later in Chapters IV and V); and (iii) the variety of components in complex stimuli and their relative capacity to constrain attentional effort. Improvements of Jones's model are the result of taking into consideration the interactions between temporal and non-temporal information and the analysis of the temporal unfolding of musical patterning; thus, the model presents a more comprehensive picture of the temporal dimension of musical structure, taking into account variations in the temporal deployment of event structure, and the ways people respond to them. People attend differently to events with high and low degrees of structural coherence, and this in turn affects time estimations. As was stated before, the ways music unfolds, and the interaction with attentional resources, allows listeners to anticipate the future course of events, including the moment at which a given temporal interval is expected to finish (Boltz, 1989; 1993).

If attention is modelled by the underlying hierarchy, we can predict that dominant, prolonged events become cognitive reference points to which attention is devoted, and subordinate, prolonging events that are associated with the former provide clues to orient expectations of continuation and/or closure, according to the linear constraints of the underlying voice-leading. We will provide experimental evidence of this hypothesis in Chapter V.

Jones's music-attending model distinguishes between two ways of dynamic attending: *analytical attending*, which is controlled by changes in the musical surface, and *global attending,* which controls attention to structural dimensions at a larger scale, specifying the general configuration of the stimulus. Both attending modalities influence the time estimation of events with low and high coherence respectively. As long as they develop in time, attending processes result in what Jones calls *future oriented attending*. Some events

are highly predictable because they contain temporal relations that interact with non-temporal information according to highly organized patterns, whilst other events are less coherent and more unpredictable because they contain less structure. Highly coherent events support a kind of 'future oriented' attention. Due to their high degree of predictability people can build and use higher-order temporal patterns, with the aim of generating expectations as to how and when they are going to finish. In the case of Western tonal music, if the notes of the underlying voice-leading are temporally organized in coherent ways, listeners will frequently anticipate not just which notes are about to come but also the temporal moment they are going to occur as a result of future oriented attention.

Thus, future oriented attention takes advantage of the global temporal structure of events, and time estimations will be the result of the confirmation or the violation of expectations about the expiration time of events. Analytical attending, on the other hand, occurs when less coherent events are being experienced. They have low temporal predictability and therefore people are not able to anticipate their future development. Instead, they are forced to attend locally to adjacent elements with the intention of organizing non-structured information. At the same time, interaction with structure prompts *attention to different parts of such structure, while others are ignored.* Jones claims that attention is flexibly selective and that, as a consequence, there is probably not a correct and stable representation of an object in the moment-to-moment process of attending.

Summarizing, in Jones's theory time is understood as a relative dimension. People can use the dynamic structure of events to attend to it in different ways. She proposes that those events that have characteristic temporal transformations are easier to anticipate in time. Structures of high temporal coherence are hierarchical and different from temporally

incoherent or non-hierarchical ones. In the temporal nested hierarchy the most prominent temporal levels are markers of the most salient non-temporal information. Dynamic attending is organized around two features: the identification of non-temporal markers corresponding to different time levels, and the organization of the temporal nesting. Jones refers to hierarchy as a temporal structure in which the temporal distribution of markers reveals nested temporal levels that are consistently related to each other at a pre-determined level by certain proportions. Beginnings and endings of temporal periods are intrinsically marked by structural changes. However, their relative salience is dependent on the context.

As in Jackendoff's proposal, the particular arrangement of the attentional mode presented by Jones unveils assumptions about the mental representation of the structure of an object (the piece of music) in time and space. As was noted before, it is assumed in this thesis that this framework may provide strong clues to study prolongational aspects of the musical piece during the experience of attending to music.

CHAPTER III

PROLONGATION AS EXPERIENCED: SOME PRECURSORS

A body of research has collected behavioural and ontogenetic evidence that accounts for the psychological reality of underlying musical hierarchy. These studies are considered as precursors to the investigation of the psychological status of prolongational structure in music. There follows a summary of this experimental work.

III.1 Abstraction of hierarchical levels

The pioneering experimental work on the assessment of the cognitive reality of the underlying structure are the seminal studies of Serafine (1988), and Serafine, Glassman & Overbeeke (1989). According to their view, one of the processes involved in music cognition is the abstraction of hierarchical levels. It is characterized as a non-temporal cognitive process, which involves formal, logical and abstract operations that are applied to music listening at deeper levels than the note-to-note level of the musical surface. In a number of experiments with children and adults, Serafine collected behavioural and ontogenetic evidence with the aim of investigating the degree to which hierarchic structure could be said to have reality or to be used in the perception of music. Experimental techniques employed involved participants performing matching and rating tasks, similarity judgment tasks and recognition tasks under the assumptions that they represented instances in which subjects' understanding of hierarchical structure was evidenced.

Overall, they found that this process is configured in human cognition at about eight years old, and that deep musical structure, derived from the processing of prolongational

components, is recognised in matching tasks, and used in similarity judgements between melodies. In particular, the main group of experiments had the aim of investigating the ability of subjects to match renditions of the hierarchic structures of musical fragments, and to compare pairs of musical fragments with the same structure but with different surface harmonies. In the first group of experiments that involved matching tasks subjects listened to three fragments of music: a model melody, a foreground and/or a middleground reduction and a foil reduction of the model melody; they were presented in a repeated listening sequence in different orders. Subjects were required to say which of the two reductions (real or foil) sounded the most like the model and also to point out the place in the repeated listening sequence at which they had made a confident decision. The researchers found that the participants were able to identify the correct structural reductions with modest success, although the success rates were quite variable across items.

An analysis of the structural quality of the composed reductions showed subjects' sensitivity to metric accent as a determinant of structure identification: subjects regarded as structurally related with the model those reductions that had the focal tones in strong metric positions. It was also found that repeating listening seemed not to make identification more likely: overall, participants were sure of their responses in the middle of the repeated sequence; however, those who were undecided did not improve with repeated hearing. Moreover there were non-significant differences between musicians and non-musicians, although musicians gave accurate responses for both foreground and middleground reductions whilst non-musicians succeeded only with foreground reductions.

In a second experiment, the aim was to investigate the relationship between hierarchic structure and aesthetic goodness. Participants listened only to the real and foil reductions of the first experiment and gave an aesthetic preference response. A correlation between the subjects' ability to identify structural reductions and their aesthetic preferences showed that they were independent. Again, there were non-significant differences between musicians and non-musicians in their aesthetic preferences.

In the third experiment the authors tried to investigate whether the modest identification ability found in the first experiment could be enhanced providing subjects with self-controlled and unlimited hearing of the stimuli, with a motivational reward when identifications were correct, and making foils 'look' more different than the correct reductions by means of the introduction of more different tones. The aim was to make the correct reductions easier to identify. A matching task and a preference task were used in this experiment. In spite of the changes introduced in the experimental task and materials the results were similar to those of the first experiment. The difference was that in this experiment the preference ratings correlated with the identification of the structure: the items whose reductions were judged more aesthetically pleasing were also those in which participants had matched the reduction to the musical excerpt more frequently.

In the general discussion of this group of experiments Serafine *et al* identified a potential problem that might have emerged from the reductive process. If subjects mentally reduced music to its underlying structure then they might also have reduced the reductions. And, given that reductions and foils were composed in such a way to be almost identical in structure but different in surface, then as a consequence of the reductive process the differences would be 'reduce away'. For this reason in the second group of experiments

the authors did not use reductions and worked instead with pairs of musical fragments that were treated with respect to their differences between surface and structure.

In the first experiment of the second group they manipulated the deep structure and the surface harmony of musical pieces in varied ways in order to form pairs with the same deep structure and different surface harmony in one case, and, conversely, with the same surface harmony and different deep structures in the other. Subjects were required to perform two similarity judgment tasks separated in time: a matching task with unlimited listening, and a rating task with immediate judgment of pairs of fragments. The results showed that subjects appeared to attend to the surface level in the matching task and to the underlying structure in the rating task. Two issues were considered as a consequence: (i) it was not clear if structural priority and/or surface priority was due to the operation of a model of music perception that begins at the deep level on the first hearing and progresses to the surface levels with repeated hearing, or if it was the result of the required experimental tasks. And (ii) the results might be due to the influence of the type of task, that is to say, a matching task might elicit a surface-like mode of perception and a rating task might elicit a judgement based on structure.

Consequently a new experiment was run using only the rating task and a number of controlled repetitions, with the aim of replicating the results obtained in the rating task of the previous experiment. Positive results were obtained with the replication of the rating task, and the distinction between surface and structure was clear.

The last experiment had the purpose of exploring if involving subjects in the performance of a different task, in particular one that involved the use of long-term memory of the underlying structure, could elicit similar distinctions between surface and structural features. In the previous experiments the rating task had required the participants use of

short-term memory. Serafine *et al* hypothesized that fragments with the same deep structure would be more likely to be 'mistakenly' recognized than fragments with different structure. Using identical materials to those used in the previous experiment, they required subjects to listen to the fragments and in the temporal interval that followed the fragment's presentation, they were required to 'rehearse' each fragment mentally. Afterwards a recognition test was performed in which subjects had to provide a yes/no response indicating if they had heard exactly each fragment before, and to use a 3 point scale to indicate how confident they were of their judgement. Results indicated that subjects accurately recognized the fragments that they had and had not heard before. But they did not support the hypothesis about the likelihood of mistaken recognition based on the sharing of the deep structure.

Taken together, the second group of experiments provided evidence of the differential recognition between surface and structure, but there were still complications related to the use of the similarity judgment in matching and rating tasks, and the use of repeated hearing. These problems highlight the question of the kind of attention (subconscious or focused) which is involved in the perception of the underlying structure. The results are not quite clear, and await additional research. The second question concerns whether it is structure or surface that remains more potent in memory. According to the results this question is wholly unknown. So, the type of attention and the type of perception which are most capable of grasping the underlying structure and the role of deep structure in memory representation needs still to be addressed.

Based on the work of Serafine et al. (1989), Dibben (1994) tested the time-span reductional component of Lerdahl & Jackendoff's *GTTM* (1983) with the aim of investigating the extent to which it represents the hierarchical structure that listeners infer

from the musical surface. The main difference between Dibben's experiments and the previous work is that she employed a more constrained version of the reductional procedure in order to avoid the complexities that emerge from the treatment of linear aspects of the prolongational theory. For this reason she used the time-span reductional analysis of *GTTM* (see Chapter II, section II.2.2.2 for an explanation of the difficulties of *GTTM* to provide a strict hierarchical account of linear aspects of the voice-leading arrangement).

She expanded the scope of Serafine et al. materials using longer musical fragments; the number of changes between foils and correct reductions were also controlled and reduced to only one per foil; she also derived 'more reduced' reductions than the reductions used by Serafine et al. from the musical fragments, in order to explore levels closer to the background, and finally she considered critically the instructions Serafine gave to the subjects; as a consequence, she changed them on the assumption that they might misguide subjects to pay more attention to surface details than to deeper levels of the underlying hierarchy. (We saw above that attention to surface details, where sensitivity to the underlying hierarchy was expected, was one of the main problems Serafine faced in the series of experiments conducted with her colleagues.) For this reason, Dibben instead used the idea of a reduction as a simplification of the model.

Dibben's experiment was a modified version of that of Serafine et al. After performing a time-span reductional analysis over the musical fragment, an analytically correct reduction was obtained. Based on the latter, five foils were created by replacing only one event in the correct reduction. The foils varied in the strength of the substitution performed. The degree of strength was measured according to the difference in hierarchy between the event that was changed and the event that substituted it, as proposed in Lerdahl &

Jackendoff's time-span reductional analysis. Three combinations of reduction and foils were produced, manipulating the degree of disruption from the model: pair 1, a reduction and a foil with greatest strength of substitution; pair 2, two foils of different strengths of substitution and pair 3, two foils of similar strengths of substitution. It was expected that the pattern of similarity judgements would be dependent on the degree to which foils were judged to be similar to the musical fragments on the base of the strength of the foil event. Thus, the more disruptive the foil, the poorer the match between reduction and musical fragment.

The experimental task required participants to listen to the musical fragment and then to one of the pairs (1, 2 or 3), and to judge which of the two members (reduction and/or foil) best matched the musical fragment. It was expected that pairs 1 and 2 should have the more correct matching responses and that pair 3 obtained the least correct matching response. The results confirmed the prediction. The ability to reject a foil as a match with the original was strongest when the difference in hierarchy between the original and the foil event was greatest. Similarly, the correct matchings were weakest in the cases of foils that replaced low-level events with other low level events.

A second experiment was run to discard an implicit effect of grammatical coherence of the foil per se on the similarity judgements. Subjects listened to the foils and had to indicate how reasonable they judged to be a given reduction of the musical fragment from which it was derived using a 5-point scale.

Differences in the perceived coherence were correlated with the results of the previous experiments and a relationship was not found between the estimated coherence of a foil and whether it was chosen as themore correct match. In summary, evidence of the use of a

hierarchical representation in the terms proposed by the time-span reductional theory of Lerdahl & Jackendoff's theory was found.

III.2 Categorization of the underlying structure by familiarity

Bigand (1990, 1994) also tested reductional representations of musical materials treated according to the time-span reduction and the prolongational reduction principles derived from *GTTM* (Lerdahl & Jackendoff, 1983) in an experiment that involved participants in assessing the degree of melodic resemblance in families of melodies. Two different structures, the 'reduced structure' (derived from the time-span reduction procedure) and the 'prolongational structure' (a kind of harmonic analysis of tensions and relaxations derived from *GTTM*'s prolongational reduction procedure) were hypothesised as corresponding to two different degrees of abstraction in music representation. Therefore, the experiment aimed at exploring the availability of those two forms of structural representation during listening tasks.

Two families of melodies were used as stimuli: family 1 had melodies with different rhythmic-melodic contours and the same harmonic structure; family 2 was created using the same rhythmic-melodic contours of family 1 but assigning different harmonic structures to them. Musicians and non-musicians were tested using a categorisation task for family resemblance, in which they listened in the first place to the group of melodies that belonged to family 1 and afterwards to these melodies interspersed with the melodies of family 2. Subjects had to estimate how well the melodies belonged to family 1.

A false family group was used as a second experimental condition: it included melodies with different harmonic structures in the presentation of the family of melodies. It was expected that there would be fewer errors in the true condition than in the false condition

if the underlying structure was perceived. Concerning the factor of musical experience, it was expected that the effect of family resemblance would be greater for musicians than for non-musicians. A third factor accounted for the two forms of underlying representation that were explored: in condition 1 all the melodies were presented in the same key and in condition 2 they were presented in different keys in order to isolate the prolongational component from the reductional component. Results indicated that both musicians and non-musicians internalised a form of underlying representation that they used to identify the family resemblance between groups of melodies. When a true family was presented they made fewer mistakes.

Concerning musical experience, it was found that musicians exhibited a higher global accuracy than non-musicians in abstracting the underlying structures. Even though musicians did better than non-musicians, there was no interaction between the factors of family resemblance and musical experience, which suggests that the effect of family resemblance was identical for both populations. It seemed as if both groups used the same listening and analytical strategies to process the musical information.

Even though the results of this experiment seem to account for a capacity of listeners to abstract structural components of tonal music, the analytical categories used by Bigand to describe prolongation do not seem to match the definition of prolongation that is widely agreed in the field of music theory (see chapter II for a thorough analysis of the concept of prolongation). On the other hand, the reductional procedures used in the experiment to treat the melodic surfaces are not conceived exactly in the terms proposed by Lerdahl and Jackendoff in *GTTM*. Bigand's interpretation of the time-span reduction solution appears to assign an absolute rather than a relative status to the pitch events that result from the application of that procedure. For this reason he posits that if the melodies are transposed

to other keys the listeners will no longer assign those notes properly to the transposed melodies. This intepretation seems to ignore the *relative component* that is the foundation of the tonal music organization since Riemannian times so far in music theory, and that, of course, is also aknowledged in *GTTM*. Finally, the procedure that he followed to derive what he calls prolongational structure is not exactly what results from the application of *GTTM*'s prolongational reduction procedure: it resembles more a standard harmonic analysis than a linear-harmonic derivation from the underlying structure of the musical excerpt. In summary, even though Bigand's subjects seem to operate with some kind of structural components when they assess family resemblance, they are not the prolongational structures that are the object of study in this thesis.

III.3 Our own previous investigation

III.3.1 Abstraction of hierarchical tonal structure

Based on the behavioural evidence obtained by Deutsch & Feroe (1981) on listeners' use of hierarchical tonal alphabets to remember melodies, and also in the voice-leading component of Schenkerian theory, Martínez (1997) collected behavioural evidence of the use of tonal information in the terms proposed by both Deutsch & Feroe's and Schenker's models. She involved participants in a performance task in which they were asked to write down tonal melodies, and found evidence of subjects' processing of the underlying structure: transcription errors were not found in the structural tones that formed the *Urlinie* of the melody (see Schenker, [1935] - 1979). It seems as if the underlying voice-leading acted as a frame that supported the processing of the remaining melodic information. Although the listener may be "lost" at certain moments during the process of writing-

down, probably due to memory "accidents" in the processing of the melodic surface level,

he or she "mentally retains" the structural path (see Cadwallader & Gagné, 1998).

III.3.2 Categorization of the underlying voice-leading by similarity

Following Serafine and Dibben's findings, my colleague and I pursued the study of the

cognitive reality of reductional representations in terms of some of the assumptions of

Schenker's theory with the aim of obtaining behavioural and ontogenetic evidence of its

cognitive representation. The experience of the underlying structure was tested in a series

of experiments that involved the activation of two cognitive processes: categorization and

similarity (Martínez 2000 a, b; Martínez 2001; Martínez & Shifres 1999a, b, c, Martínez &

Shifres 2000; Shifres & Martínez, 1999; Shifres & Martínez 2000 a, b, c).

As we saw in Chapter I, according to prolongational theory, tonal coherence is the result of

relationships between the musical surface and deeper structural levels. Its experience

involves building *reductional representations.* A reductional representation is defined as a

simplification of a series of events into a single abstraction at a deeper cognitive level. It

was assumed that if reductional representations occurred during music listening, they

would be used to make comparisons between melodies. Thus, an experiment involving

similarity judgments tasks was considered appropriate to provide insights of such

underlying representation. As we saw before, both Serafine and Dibben aknowledged the

existence of methodological difficulties to create experimental conditions to study

reductional representation. Serafine tried to overcome the problem that might have affected

the subject's responses during a matching task between a melody and its rendered

reduction. In further experiments, instead of using rendered reductions that could also be

reduced during the matching task, she worked with comparisons between musical

fragments of the same structural level. Bigand also tried to manipulate surface and structural differences forming families of melodies, although as discussed above, he did not approach the underlying structure quite in the terms proposed here.

In order to understand better the methodological problem that emerges when the relationship between the surface component and the underlying component is considered, it could be useful to analyse first the problem that emerges when a standard experiment on similarity judgements is designed. Two conditions can be arranged: in condition 1 we have the original sample and the foil with the same underlying voice-leading (uvl) and different surface; in this case we can predict that similarity will be based on the uvl. In condition 2 we have the original sample and another foil with different uvl and the same surface. In this case the prediction is that similarity will be based on the surface component. But if it is possible to modify the melodic surface without modifying the uvl as in condition 1, it is not possible to satisfy the requirements of condition 2 because if the uvl is modified, the surface will be modified as well. However, it is possible to treat the melodic surfaces of the musical samples in order to operate with them as an experimental variable.

Thus, in our experiments it became necessary to control precisely the changes that occured at the level of the musical surface when the underlying structure was manipulated, so as to know if the listeners'judgments were based in the changes operated on one or the other structural level.

In the first group of experiments we looked for evidence of the reductional representation using unaccompanied melodies. The experimental task required subjects to compare a model melody (A) with two foil melodies, one with the same uvl (B) and the other with different uvl (C). The challenge was to find an experimental method that accounted for similarity judgements based on the uvl and not on the musical surface.

A systematic procedure was followed in order to control the musical surfaces during the elaboration of the musical samples. Twenty melodies from the repertoire of Western Art Music were selected. Two kinds of experimental controls were performed in the elaboration of the foils: (i) control of the compositional process. In order to do this the following sequence was followed: 1. selection of a model melody (A); 2. reductional analysis of the model melody and derivation of a structural reduction (R1); 3. substitution of some notes in R1 in

Figure III 1. Procedure followed in the composition of the stimuli. Step 1:selection of a tonal melody; step 2:reductional analysis of the surface level (R1); step 3:derivation of a second musical reduction from R1 (R2); step 4:composition of a musical surface from R2 (C); step 5: composition of a musical surface from R1 (B) that is similar to A and C. Extracted from Martinez and Shifres (1999a).

order to obtain a new reduction (R2). R1 and R2 were composed applying the voice-leading principles derived from strict counterpoint; 4. composing-out of a new musical surface (C) from R2, with similar surface and different uvl, trying to minimize the changes with respect to A; 5. composing-out of a new musical surface (B) from R1, with similar surface and same uvl, trying to minimize changes in relation to both A and B. B and C had the same number of notes and the same rhythm as that of A. The substituting notes in B and C kept diatonic coherence in relation to A (this is shown in Figure III.1). (ii) Control of the degree of theoretical similarity of the melodies. The similarity of B and C in relation to A was estimated by means of a series of indices of similarity. The number of substituting notes and their duration were calculated and an index of similarity was applied. Differences were not significant. An adapted version of the control of structural weights of Serafine et al (1989) was applied to measure the metric structural importance of the substituting notes.There were non significant differences in the metric structural importance of the substituting notes. Primacy and recency effects were also controlled under the assumption that the substituting notes that were located at the begining and at the end of the melodies could affect the similarity judgements. Differences were non significant.

The melodic contours of A, B and C were also correlated in order to obtain a similarity measure of their degree of association. The number of semitones of the intervals was counted and a positive and/or negative sign was assigned to each interval according to its ascending and/or descending direction. A Pearson correlation coeficient was run to measure the association between the numerical series thus obtained. It was hypothesized that according to the procedure that was followed and the restrictions imposed on the composition of the melodies, this measure of the contour association might represent a

measure of theoretical similarity of the melodic surfaces. The procedure followed can be seen in Figure III.2.

Two experiments were run to test the use of the uvl in the similarity judgments between pairs of melodies. In experiment 1 (Martínez & Shifres, 1999a) 190 adults, divided in three groups of musical experience, volunteered to participate. It was predicted that subjects would rate the melody which shared the uvl as the most similar to the model. In each trial subjects were required to provide a similarity judgement in a simple comparison task: they listened either to the pair A-A-B or A-A-C, and had to estimate the similarity between A and B and/or A and C using a 7 point scale.
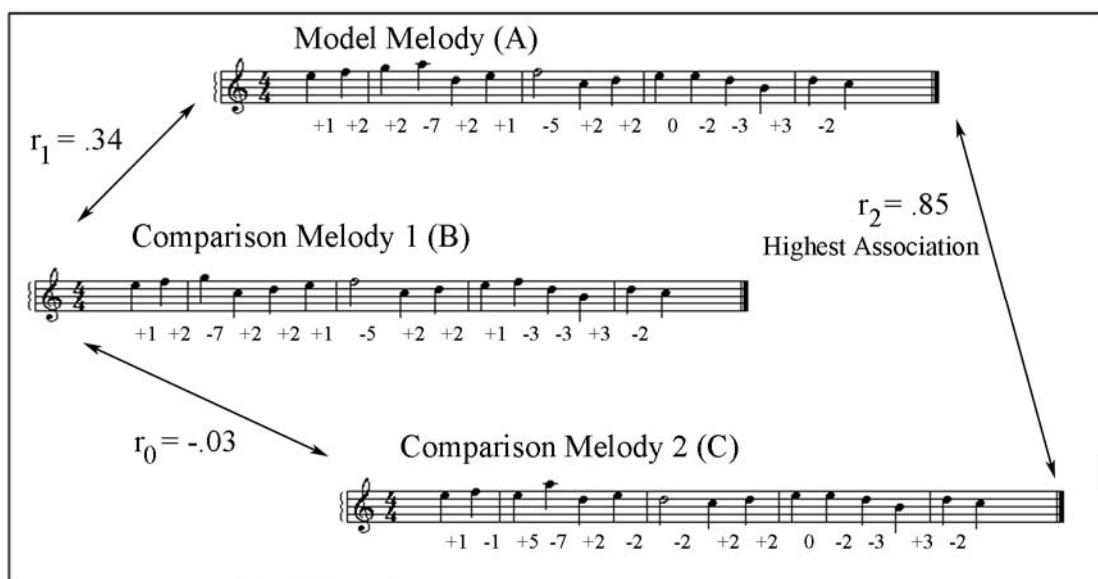
Figure III 2. Procedure followed to extract a comparative measure of the level of similarity of the melodic surfaces. r values measure the degree of association between the numerical series that results from the analysis of melodic intervals in terms of the number of semitones, and of positive and/or negative sign indicative of the interval direction. Extracted from Shifres and Martinez (1999).

Results showed a marginal significant difference of a higher similarity rating for the pair A-B, indicating that subjects seemed to make a moderate use of the uvl in their similarity

judgments. There were non-significant differences between groups in terms of their musical experience. Experiment 2 (Martínez & Shifres, 1999b) was an adapted version of experiment 1 and was administered to 774 children divided into three age groups. Results also indicated a moderate use of the uvl in the similarity judgments.

Results confirm a developmental trend that was detected in previous studies relative to the understanding of tonal organization (Serafine, 1988; Lamont and Cross, 1994). Evidence of a very poor perception of differences between surface and structure at the age of 6 was found. Competence seems to be acquired at 7-8, and increases with age. The developmental curve that was found is similar to others already reported (Serafine, 1988). However, these results require more investigation.

From the treatment applied to the musical surfaces of the samples of this experiment two groups of musical samples were formed according to the relative degree of association between the contours of A, B and C. In group 1 the highest association occurred between B and C and in group 2 the highest association occured between A and C. A detailed analysis of the responses given by the children to the musical excerpts showed that children estimated B (same uvl) as the most similar to A in group 1, while they rated C (different uvl) as the most similar to A in group 2. That is to say, in the first case, children judged similarity based more on the uvl, while in the second they judged similarity based on the contour more than on the uvl. Previous experiments on melodic contour (Dowling, 1994) had strengthened the value of this musical feature as a perceptual attribute of immediate access. The highest similarity between the contours of B and C resulted in a neutral salience of contour that could not be used to provide the similarity judgments. In these cases, children's decisions were based on the underlying structure. These findings seemed to indicate that the correlation measure used to estimate the degree of association between

the melodic surfaces could be used with a high degree of certainty as a model to predict the theoretical degree of similarity between the surface component of tonal melodies.

In summary, experiments 1 and 2 provided evidence that listeners develop sensitivity to the underlying voice-leading in similarity judgments tasks. Ontogenetic evidence of its presence as a cognitive process was detected in children of 7-8 years old (Martínez & Shifres, 1999b). This evidence confirmed former findings of Serafine (1988) regarding the acquisition and development of structural hearing as a cognitive non-temporal process. The use of structural hearing was also found in adults (Martínez & Shifres, 1999a). A distribution of the similarity judgments according to the theoretical contour similarity predicted by the correlational model (Shifres & Martinez, 1999) accounted for the predictive value of this model to treat the melodic surfaces in order to isolate the uvl as an experimental variable.The results of experiments 1 and 2 encouraged the design of a new experiment in order to go further in the experimental testing of the cognitive reality of the underlying representation in tonal melodies. A hypothesis of a theoretical rivalry between the underlying structure (understood as the uvl) and the melodic surface (measured according to the note-to-note relationships of the melodic contour) that might affect the similarity judgments was developed, inclining listeners to provide their similarity ratings according to the predicted rivalry between both components.

In experiment 3, 15 tonal melodies were selected as stimuli. This time, the comparison task required subjects to judge the similarity between trios of melodies. They were formed by a model melody (A) and two comparison melodies that were composed with the same uvl and similar surface (B) and with different uvl and similar surface (C). This time the correlational model of theoretical contour similarity (Shifres & Martínez, 1999) was used in the compositional process to monitor the degree of theoretical similarity of the samples.

Two groups of trios with different degree of relative similarity between A, B and C were formed: in group 1, the highest contour association was between melodies A and C; in group 2 the highest association was between melodies B and C. As melodies A and B had the same uvl, it was predicted that this treatment would create a rivalry condition between the surfaces that could provide indicators of the relative use of surface and /or deeper structural levels in the production of the similarity judgements.

146 adults with different levels of musical experience took part in the experiment. The experimental task consisted of listening to the sequence AB - AC (or AC - AB) and to judge which of the two comparison melodies (B or C) was the most similar to A. Subjects also had to estimate the level of confidence of their answers using a three point scale (very sure - not so sure - not sure). It was hypothesised that listeners would judge melody B as more similar, although the different degrees of association between the melodic surfaces might cause confusions in the responses. Two experimental conditions were arranged: in condition 1, the group of melodies with the highest contour association between A and C was presented; it was predicted that when the subjects judged the similarity between the pairs AB and AC, they would hesitate because the rivalry between B and C, with respect to A was high; in condition 2, the group of melodies with the highest contour association between B and C was presented; it was predicted that when the subjects judged the similarity between the pairs AB and AC they would not hesitate because the rivalry between B and C with respect to A was low and the uvl would be the most salient attribute.

Data for B/C responses and confidence ratings were translated into a single score ranging from 1 (very certain C) to 6 (very certain B), where 3 and below represented "C" and 4 and above represented "B". Results confirmed the prediction: subjects always tended to judge melody B (same uvl) as the most similar; however differences in their ratings showed that

structural and superficial attributes compete in their relative salience, causing different levels of perceptual rivalry.

Post-hoc analyses of the results of experiment 3 were performed applying alternative models that analyse hierarchical aspects of musical structure (Martínez & Shifres, 2000). The purpose of these analyses was to test the pertinence of alternative theoretical models to describe the melodic attributes that are modified when the uvl is modified. It was expected that these models might explain better the results of the experiment and/or might be useful in accounting for additional information that might help in finding a more precise estimation of the real incidence of the use of the uvl in the similarity judgements. So, the stimuli used in experiment 3 were analysed according to the following models: models of melodic contour (oscillations model: Schmuckler, 1999; combinatorial model: Queen, 1999); model of tonal weights (probe tone ratings: Krumhansl, 1990) and model of event hierarchies (melodic anchoring: Bharucha, 1996).

None of them was more successful in explaining the results of the experiment. Therefore, evidence about the parsimony of the method developed to analyze the similarity of the melodic surfaces was obtained.

Martínez (2000 a, b) and Martinez (2001) discussed the pertinence of the similarity judgement paradigm for the study of prolongational aspects of the underlying musical structure, highlighting the need to take into account the influence of contextual factors on similarity judgments. Applying the model of asymmetry (Tversky, 1977) in a post hoc analysis of the results of experiment 3, she found that the order of presentation of the melodies in the comparison pairs, the condition of subject or referent of the melody in the comparison task and relative salience of the melodies according to whether they belonged

to one or the other contour group, were contextual factors that created conditions of asymmetry that modified the strength of the similarity ratings.

According to the results of this analysis, (i) when the order of presentation of the target melody (A) and the comparison melodies (B or C) is direct and not inverse, it increases the estimated perceptual similarity; that is to say, in AB-AC, B is judged as more similar to A than in AC-AB, because the pair of higher prototypicality (AB) is heard in the first place; AB is more prototypical than AC because both melodies share more attributes; in this case they share the contours and the underlying structure; (ii) the comparison between two pairs of stimuli involves the relationship between a referent and a subject: in the group AC-AB, A is the referent and C and B are the subjects that are compared against A; the referent condition says that in the case of AC-AB, AB is going to be estimated as more similar than AC because AB is the pair that is more prototypical; (iii) this condition represents the relative salience of the melody according to the contour group to which it belongs. It predicts that the degree to which a melody belongs to a group of melodic contour increases the prototypicality of the melodies whose contours are more similar. Therefore, B and C will be perceived as more similar in the contour group BC than in the contour group AC. This prediction has guided the realization of experiment 3.

A repeated measures Anova showed a significant effect of the three factors and their interactions. According to the analytical procedure followed in this post hoc study, similarity seems to be the result of making comparisons between sets of similar and/or different attributes of the surface and of the underlying levels of the musical structure. Contextual changes in the experimental conditions elicit changes in the diagnostic value of those attributes. In our case, the interaction of the contour similarity groups, the order of presentation of the melodies in the trios, and the condition of subject and/or referent of the

melody to be compared (Tversky 1977) increased and/or decreased the diagnostic value of those attributes and determined changes in the strength of the estimated similarity, without nevertheless modifying the essential relationships between surface and underlying structure.

Furthermore, Shifres & Martínez (2000 b, c) studied the incidence of music performance in the communication of the underlying voice-leading, testing the listening experience of rendered underlying voice-leading reductions extracted from renowned commercial interpretations of a fragment of the Bourré I from the Suite no. 3 in C Major for solo cello by J. S. Bach. Participants were involved in two different tasks: a matching task and a rating task. Results indicated that listeners' answers were different according to task requirements, confirming the findings of Serafine that subjects' responses are different according to the task that is involved.

In summary, the behavioural and ontogenetic evidence provided by these experimental precursors of the experience of prolongational aspects of the underlying musical structure indicates that listeners show sensitivity to the underlying musical structure in similarity judgement tasks between pairs of melodies with same and different underlying voice-leading. Experimental control of theoretical similarity between melodic surfaces and the underlying voice-leading (Shifres & Martinez, 1999; Martínez & Shifres, 1999b) gave rise to the finding that similarity judgments are influenced by a rivalry between different hierarchical levels of musical structure (Martinez. & Shifres, 1999c; Shifres & Martinez, 2000a). Alternative models that analyse structural features of music perception could not explain this evidence better than the model used in our experiments (Martinez & Shifres, 2000). Nevertheless, the use of similarity judgement tasks as an experimental procedure to study categorization must take into account that contextual factors might be involved in

similarity measure variations (Tversky, 1977). Experimental work on the perception of similarities between pairs of musical fragments found that similarity judgments seem to be context-specific (Lamont and Dibben, 2001). In our experiments it was also found that similarity judgements are influenced by the contextual factors of relative salience and order (Martinez, 2000a; 2000b; 2001). Moreover, it was found that the nature of the cognitive task in which the subject is involved needs to be taken into account in seeking further evidence of the cognitive reality of this construct (Shifres & Martínez, 2000b).

The impact of perceptual similarity is also apparent in the structure of people's everyday life theories. When categories are organized at different levels of abstraction, the basic level seems to be the most natural psychological level on which to categorize the environmental experience of reality. Familiarity, prototypicality and exemplarity are three characteristic features that are considered descriptors of the use of categories to make sense of reality. Psychological essentialism (Medin & Ross, 1997) integrates similarity and explanation, suggesting that people act as if things had underlying essences or natures that make things as they are. This conception does not refer to the way things are, but to the way people approach them, suggesting that people adopt a heuristic essentialism, in the sense that they build hypotheses about things, based on shared deeper properties, that is to say, on similarities. It is therefore reasonable to think of similarity as a general guide or heuristic to perform a categorization activity. It seems that human perceptual and conceptual systems have evolved in such a way that the frequent and correct use of the essentialist heuristic occurs as a consequence of the adaptative behaviour. This observation suggests that the use of similarity judgements in order to make classifications is effective, and may be more related to knowledge of deeper principles or attributes than to sensitivity to surface characteristics (for theoretical developments that discuss the sufficiency of the

similarity/categorization model to provide an explanation of the way music is understood in cognition see, for example, Ockelford, 2004).

III.4 Prolongation as experienced: summary and prospects

In this chapter a review of the precursors in the experimental study of prolongation has been presented. Aspects of two of the main theories of hierarhical representation, Schenkerian theory and Lerdahl and Jackendoff's *GTTM* have been subjected to experimental testing. Overall, it was found that the underlying hierarchy has a status of cognitive reality in tonal music representation.

According to what was underlined in chapter II, a concept or an idea gains cognitive reality as long as different types of converging evidence are obtained by means of scientific investigation. Out of the four kinds of evidence that Cohen (2000) identifies as accounting for the cognitive reality of a given knowledge, in our own previous research we have collected behavioural and ontogenetic evidence of the representation of underlying levels of tonal music. In particular, similarity judgement measures have been used to study the mental representation of the underlying structure of tonal melodies, and it was found that these measures of typicality are useful to account for the subject's categorization of melodic surfaces in terms of an underlying simplicity. The properties that emerge from the underlying structure appear to elicit the subject's activation of an optimal level of representation in order to produce the similarity judgment response.

Besides, in our studies on similarity, ontogenetic evidence was found of the acquisition of structural hearing that confirms previous findings about the presence of this non-temporal process in human cognition at about seven years old. Perceptual melodic similarity involves the recognition of a structural identity between stimuli on the basis of a common

attribute, the underlying structure, and a different attribute, the surface. Overall, similarity is defined as a multivariate function of the differences between pairs of attributes that objects posses (Cambouropoulos, 1997). Usually, similarity is considered as a kind of partial identity, given that two objects are similar if they share some but not necessarily all their properties. According to our results, it is estimated that the cognitive representation of an underlying structure, an attribute that reduces the musical surface, is possible when *decentration* develops in cognition. Decentration is defined by Piaget as a cognitive process that allows subjects to take into account all the attributes of the perceived object and to recognize their interdependence, instead of basing the estimation in only one of them in particular. Experimental results show the presence of structural hearing at the age in which concrete operations begins in cognition. As pointed out in Chapter II, developmental evidence from observational studies of the representation of hierarchical organizations shows that some hierarchical levels are acquired before others and also that some hierarchical levels tend to persist through time. We know from previous research (Dowling & Harwood, 1986) that melodic contour appears in development very early in cognition. The results of our experiments on similarity judgements with children show that, once sensitivity to underlying structure is acquired, it persists and increases during childhood.

Concerning the cognitive reality of hierarchical organizations it is important to underline that when cognitive hierarchies are investigated the evidence may be subject to alternative interpretations, and the obtained results might also be explained without invoking those hierarchies. In our studies of similarity we contrasted the predictive value of the theoretical measures provided by the correlational model to control the degree of similarity of melodic surfaces with other measures provided by alternative theoretical models of tonal hierarchy,

and found that the most parsimonious explanation of the behavioural evidence obtained in our experiments was provided by the correlational model.

The experimental testing of the hypothesis of perceptual rivalry between the attributes of the melodic surface and those of the underlying structure showed that results exhibit differences according to the structural priority of the hierarchical level that is taken into account to produce the response. Thus, it was possible to find that listeners are sensitive to those two different levels and that perceptual rivalry between them elicited subjects to focus more on one than on the other component.

Other indicator of behavioural evidence when the hierarchical levels of structural organizations are investigated can be found in the differences of error rates in the subjects' responses. In a previous study on memory for melodies, Martínez (1997) found more error notes in the transcription of lower level embellishment tones whilst fewer error notes were found in those structural notes that belonged to the underlying voice-leading structure of the musical samples. This result seems to indicate that higher-level structural representations are more stable and remain in memory.

However, our previous investigations warned about the influence of contextual conditions that bias similarity judgements, a problem that had been previously identified in the similarity research in other knowledge domains (see Chapter II). Furthermore, when subjects were involved in a comparison between melodic fragments and rendered reductions of those fragments that required the performance of a matching task and a rating task (Shifres & Martínez (2000 b) an effect of the task in the production of the responses was found.

However, in spite of the compelling evidence of the cognitive status of underlying representation in tonal music provided by the previous studies, the shortcomings identified indicate that in order to go further in its investigation it is necessary to broaden the scope of this inquiry, and explore in more detail additional instances in which the subject's sensitivity to prolongational structures could be elicited, using other cognitive processes that involve participants in different experimental tasks.

If the aim is to go further in the search for evidence of the cognitive status of hierarchical representations in tonal music it is also necessary to test other analytical dimensions that previous studies did not tackle.

Therefore, in this thesis, other cognitive dimensions potentially involved in the experience of the underlying hierarchy of tonal music are pursued. The first part explores dynamic aspects of the process of attention to music hierarchy. The type of control processing that takes place when attention to the prolongational structure is involved is investigated. The extent to which the focus of attention changes according to the structural demands of music fthe assumption that prolongation functions a constituent structure. Reaction time responses, which seem to provide good behavioural indices of the hierarchical status of segments in constituent representations are used in the experiments to be pursued. It is expected that these measures account for the listeners' sensitivity to boundary locations, providing behavioural evidence of the cognitive representation of the prolongational structure as a constituent hierarchy. This approach is theoretically developed in Chapter IV and experimental work is reported in Chapter V.

The second part of the thesis focusses on alternative views in the domain of human cognition with the aim of looking for answers to the cognitive status of prolongation as experienced in the terms proposed by Schenkerian theory. This search is based on the

belief that the accounts of prolongational theory provided by the view of the computational mind offers only a partial answer to the problem of its cognitive representation. There is an idealistic and imaginative component that is implicit in the ontology of the Schenkerian view that the cognitive science approach does not tackle. If the aim is to go further in the investigation of the psychological status of prolongation it is necessary to explore alternative approaches to human cognition. It will be hypothesized that *image-schemas* - which are a particular type of mental structure of embodied knowledge that operates at the basic level of human experience - shape the metaphorical processes that take place during the experience of prolongation. The ways in which these image-schematic structures operate might reveal the kind of mental dynamic organization that subjects put into action in order to use embodied knowledge to understand the musical domain. As was posited in Chapter II, some models that were created as metaphors or imaginative constructs become accepted as evidence of their cognitive reality is found. Thus, it is the aim of the second part of this thesis to provide some evidence of the psychological reality of this construct as a cognitive metaphor.

This approach will be theoretically developed in Chapter VI and experimental work will be reported in Chapter VII.

CHAPTER IV

EXPERIENCING PROLONGATION AS A LINGUISTIC CONSTITUENT

IV.1 Modelling syntax: constituency and dependency as two structural components in tonal music

In this chapter a hypothesis about the syntactic nature of prolongation in tonal music will be posited and its implications for its structural status will illuminate the experimental work that will be presented in the following chapter. I propose that prolongation could be assigned a syntactic status that somehow shapes sensitivity to the underlying structure during music attending.

As stated in chapter I, understanding musical structure as hierarchical highlights the idea that (i) some events are structurally more important than others; (ii) events can be related over spans of different sizes and (iii) musical surfaces are derived from the elaboration of structurally significant events by other ornamental events that confer upon them the peculiar characteristic designs we find in the countless examples of tonal pieces.

Application of a syntactic approach to the study of music hierarchy is not new. It owes its existence to the widespread influence exerted by linguistic theory on music theory and music cognition (see Chapter I). Recent research in the field of neuroscience proposes a specific point of convergence between syntactic processing in language and music (Patel, 2003).

Syntactic theory describes the structure of complex objects, formalizing a view of structural relations between the units of a specific domain and describing the ways in which such units are organized to form them. The term 'complex' characterizes a description of an object in the following terms: i) the object is divisible to parts called constituents; ii) there are different types or categories of parts; iii) different parts are organized in specific ways and iv) each part plays a specific role in the structure as a whole (Horton, 1999; 2001). Thus, syntactic descriptions of aspects of tonal organization inspired by linguistic models provide useful insights about the cognition of music hierarchy and prompt a hypothesis concerning the experience of prolongational aspects of musical compositions based on linguistic concerns.

In the domain of syntax, a description of a syntactic organization acknowledges the fact that its structure can be understood in compositional terms. The term 'compositional' will be used in this chapter in the sense that it has within linguistics. The principle of compositionality states that the meaning and functionality of a syntactic unit are seen as derivable from those of the smaller units of which it is composed and the constructions in which they stand. For example, *beautiful girls* is a syntactic construction in which 'girls' is the head of this constituent unit and 'beautiful' is the modifier. In the domain of musical syntax, a syntactic description of a musical piece can also be understood by analogy in compositional terms. The piece is viewed as an object that can be divided into significant units that, as they unfold in time, are subsumed themselves into increasingly longer streams of events. Understanding a tonal piece in compositional terms implies the idea that its musical content can be explained as a function of the content of its components. In compositional analysis, tonal structures result from the combination of events according to specific properties, in which determine its syntactical content.

The most usual graphic representation of the syntactic system above described is that of a tree diagram. In tree graphics, the hierarchical structure of an object exhibits two characteristic features: precedence and dominance (Horton, 2003). Precedence deals with the temporal nature of music, requiring that some events appear before others; precedence displays the left to right ordering in the concatenation of successive constituents and defines the linear structure of the object. Dominance, on the other hand, concerns the hierarchical grouping of the constituent structure; in order for discrete events to be integrated into more comprehensive syntactical units, some of them must govern the operation of others. This kind of hierarchical relation is usually represented by displaying the correspondent ordering of nodes in the hierarchy from the root at the top to the terminal nodes at the bottom. In tree representations those events located at the lowest level of the tree structure are called terminal nodes: they dominate no other events. Non-terminal nodes, on the other hand, represent strings of elements, where some of them dominate others. The property of a node to dominate over others is called projectivity. A tree diagram is shown in Figure IV.1.
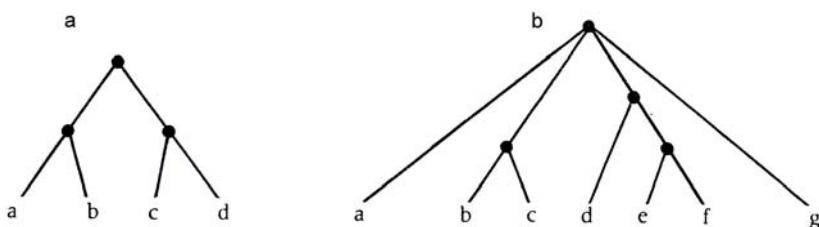
Figure IV.1. Schematic tree diagrams representations of hierarchical organizations. Terminal nodes are named with letters. Non-terminal nodes are highlighted in black dots. (Extracted from Cohn and Dempster 1992, p. 159).

Structurally speaking, the architectural hierarchy of a tonal piece concerns the recursive application of the dominance relationship. As was seen before (see Chapter I) there are,

however, two conflicting issues around the concept of dominance that have not been fully

elucidated so far: the first of them refers to qualification of the status of dominance in

terms of a representational or an inclusional hierarchy (Cohn & Dempster, 1992) and the

second concerns the scope of recursivity: that is, whether recursivity applies to the whole

underlying hierarchy - as is the case is in 'strong' science systems approach - or if it works

well within levels but is not fully applicable in the same way between levels (Schenker

[1935] - 1979; Horton, 1999) (see in Chapter I a detailed account of this issue). This idea

requires facing the problem of considering a definition of music either as a well-structured

or as an ill-structured domain (Spiro, 1987 in Covington and Lord, 1994). We will come

back later to the discussion of this issue. But by now, it can be said that the cognitive

paradigm of strict recursivity in hierarchical organizations represents - from the point of

view of theoretical modeling - just the *desire* of a musical mind to behave in this way. In

order to test a hypothesis of strong recursivity it would be necessary for an experiment on

music cognition to demonstrate a capacity to provide a full behavioural picture at the

different levels of the underlying hierarchy all at a time. Experiments using a continuous

response paradigm, that could have been potentially useful to provide a more

comprehensive picture of this kind of mental operations, have been the subject of strong

criticism (Schubert, 2001). Different experiments can offer partial experimental probes of

the cognitive experience of hierarchy at determined levels, but they cannot offer the whole

picture at once. Therefore, whether recursivity is fully applicable throughout the whole

underlying hierarchy or if the idea of reduction applies well within each hierarchical level

is a problem that remains.

A syntactic analysis of tonal structures can be performed which takes into consideration

the operation of two types of systems: constituency and dependency. In what follows both

systems and their mutual relations are going to be explained and afterwards some epistemic problems related to the status of prolongation in the terms offered by these two syntactic systems are going to be presented.

IV.1.1 Constituency

A constituent is a structural unit that is formed by events that group together on the basis of sharing a structural characteristic. Proceeding by juxtaposition, adjacent elements are combined with each other in order to form concatenated streams of events. As a consequence of the grouping process, constituents are subsumed into larger components, and the compositional process continues recursively until the whole structure of an object is obtained (see the previous section for concerns about the idea of recursivity as applied to hierarchy in music). The nature of constituency is then combinatorial and generates a hierarchy of part-whole relations.

Constituents may be of different sizes, the smallest of them being those syntactic formatives to which an independent function can be attributed. In tonal music, for example, meaningful units can range from a single event, such as a scale-step, or groups of successive and/or simultaneous events, such as chords or musical motives, to the whole piece of music.

Whether they are single elements or groups of elements, constituents are self-contained units. They have distinct features that contribute to their definition: in a given passage, for example, constituents may "perform" coherent structural roles, such as a suspension with its resolution, or a cadence, among others.

Given the regularities of tonal organization, constituents have characteristic distributions through the musical piece. As a result of this peculiarity, substitutions between constituents

that belong to the same syntactic category are possible, even if they are not of the same size. Thus, a minimal event with the same characteristic distribution may replace a group of events: for example, a sequence of chords in a given passage can be replaced with a single chord or an equivalent construction without changing the overall tonal structure of that passage. In even more sophisticated instances of musical elaboration, for example when a chord is subjected to changes by means of embellishment procedures, the replacement figures are entities that are realized as different surface configurations, while keeping the underlying structural identity of the passage.

On the other hand, syntactic constituents and their characteristic distributions are constrained by the musical contexts in which they occur. That is to say, the same note, chord, or motive may convey different meanings in different musical settings.

The way a constituent operates in a given structural context depends not only on the constituent's own category but also on the other constituents' categories it enters into combination with; that is to say, on the resulting arrangement between them. Different kinds of arrangement result from the linear orders produced by the concatenation of constituents, and also as a consequence of their hierarchical grouping.

If the concern is to analyse constituency as a cognitive experience it is necessary to place this structural feature in the ongoing musical context; in order to do so we need to relate it to the temporal development of music. The unfolding of a musical composition provides strong clues relative to the way its musical surface can be parsed into discrete constituents, which, in turn, are recomposed into increasingly larger syntactic wholes as they occur in time. Self-containment and syntactic identity are important structural qualities that, among others, facilitate parsing because they help frame the constituents' *boundaries*. The contexts in which constituents occur also provide important structural information that

facilitates identification of their boundaries at different levels of the hierarchical structure of the piece.

IV.1.2 Dependency

Dependency, on the other hand, characterizes an object in terms of the structural priority of its components. Accordingly, dependency relations account for the fact that some elements are subordinated to other superordinate elements on the basis of their structural qualities, the latter in turn being recursively subordinated to other superordinate elements until the whole hierarchical organization is obtained (see Chapter I for more information about tree hierarchies). Structurally speaking, dependency is therefore that aspect of an object's organization that exhibits a hierarchy of relations between leading and subordinate components throughout the whole network of events.

Analyzing the internal form of a constituent unit, it is possible to identify in the group of events one of them, called the constituent's *head*, which is more important than the rest of them, being them subordinated to it. Dependency is therefore a property of the syntactic structure manifested by means of the relationship between the head and the subordinate elements of the constituent unit.

In order for a tonal event to have a structural role that enables it to *perform* as a constituent's head, it must be describable as a distinct representative of some scale-step. In order to fulfil this condition, it must be consonant at the structural level at which its operation is being considered. Subordinate events, in turn, are dependent on the basis of the elaborative roles played by them with respect to the constituent's head. The musical effect of the relationship head/dependents is that the latter *prolongs* the head, and this is essentially why a constituent as a whole projects the structural quality of its head. The

property of the constituent formants - the events that form a constituent unit - to extend the syntactic category of the head, contributes to the constituent's syntactic definition. A constituent is therefore the expansion of its head, and the head imbues the unit as a whole with its character at a more abstract level of the hierarchical organization of the piece. To the extent that the structural properties of the constituent's head maps onto the entire constituent, or in other words, to the extent that they are distributed on the constituent as a whole, it is said that a constituent is the *projection* of its head.

All that was noted above about the constituent organization of tonal music and the particular relationship between the constituent's head and the constituent's formants elicits a description of the constituent's operation according to the terms provided by the prolongational theory. Hypothesising prolongation in terms of constituency implies assigning the group of events that form a constituent the property of extending the influence of one of them, the prolonged event, over the rest of the components, the prolonging events. Moreover, instances of the voice-leading arrangement of a musical composition can be found in the underlying linear patterns that are generated as a consequence of the connections produced between the heads of the concatenated constituent units.

Based essentially on the capacity of the subordinate events to depend on the constituent's head, prolongational derivations extend the prolonged event's *presence* along the syntactic unit. The scope of the head is thus expanded in such a way to imbue the unit as a whole with its character, at a more abstract level of the hierarchical organization. By means of this liaison, those elaborative processes that occur in the constituent unit can be understood as rendering different surface *readings* of the syntactic components of tonal structure. It is a psychological assumption that the constituent's head is somehow *mentally retained* all

along the constituent unit through the prolonging events that extend it (see in Chapter I a comment on the concept of mental retention).

IV.1.3 The relationship between constituency and dependency

Even though constituency and dependency systems are highly related, the type of information provided by each of them is not identical. In constituency systems a unit is divided into other units; in dependency systems a unit governs other units.

A tree representation of a constituent structure denotes part-whole relations and shows how some parts are included in others; a tree representation of a dependency structure, on the other hand, denotes a head-subordinate relation, that is to say, how one event governs others. An example of a dependency representation is shown in Figure IV.2.

A dependency representation provides information that cannot be derived from a constituency representation. In the former, it is possible to account for a description of relationships between heads and dependents, while the latter concerns the way elements group together in increasingly larger streams of components. Therefore, it is important to take both into account in order to provide a full description of the syntactic structure of a piece of music. It is also important to notice that although any type of dependency structure is a constituent, the other way around is not necessarily true. There are constituent structures that belong to the same functional category, for example, those coordinated structures formed as a consequence of repetition; in this case it is not possible for us to construct a mental representation based on dependency because we tend to understand them as being of equal significance (examples of this category will be found later in the experimental work presented in Chapter V).

The comprehensive nature of the constituent organization provides a strong reason to

hypothesize our understanding of the syntactic structure of tonal music through

constituent-structure representations.

Figure IV.2 Dependency analysis - in letters and arrows - of the first phrase of the Bach chorale 'O Haupt voll Blut und Wunden' (below). It was derived from the reductional representation (adapted from Lerdahl and Jackendoff 1983, p.115) presented both in tree notation and traditional notation above.

"But note how we build such larger structures: not by adding in a single event at a time, but

making the piece itself one constituent part of a larger structure which then enters into

construction with the other constituent parts of that structure; or perhaps by expanding a

particular constituent within the original piece by embedding other constituents within it. In

other words, we understand tonal pieces to be composed (in the technical sense) not from

individual events, but from syntactic parts, that is, constituents" (Horton, 2003, ch. 3, p. 47).

Thus, the combinatorial and grouping configuration of constituency indicates that music

appears to be understood more in terms of structural roles fulfilled by groups of events

rather than just individual events.

"Of course [   ] it is the head that gives the constituent its special functional character; the

subordinate events within the constituent serve to elaborate the head. But the point here is that

it is the entire constituent that performs the structural role in question, rather than just the head"

(Horton, 2003, ch. 3, p. 48).

In summary, both constituency - the concatenated organization of part-whole relationships

between groups of events - and dependency - the hierarchy of relations of dominance-

subordination between events - are necessary to account for our encoding of tonal relations

in music.

IV.2 Prolongational status of the syntactic structure

The notion of dependency highlights the hypothesis that musical events in tonal pieces are

experienced in terms of some kind of hierarchy of structural priority. As was highlighted in

Chapter I, consideration of the nature of this priority poses an epistemological problem for

the study of the status of reality of the underlying musical structure. It is not clear for

example, if dependency is informed by an inherent property of pitch events that determines

per se the hierarchical structure of tonal pieces. Although the understanding of certain

features of tonal organization, such as the scale-degree identity, have been found to correspond to a ranking of events that play a significant role in shaping the way tonal music is experienced, dependency, as a feature of musical syntax, might shape our experience in a rather different way (see for example comments on Krumhansl's tonal hierarchy and Bharucha's event hierarchy in Chapter II). According to this, dependency relations would perform significant roles in contextual frames in which structural priority emerges as a result more of the compositional nature of the tonal organization than due to some abstract inherent stability of pitch events.

Providing coherent explanations about the ways events link to each other and also showing which events elaborate which other events, dependency relations seem to be more the result of the structural roles played by tonal events in different contexts than to owe their status just to the syntactic categories of the events. This issue was highlighted by some music analysts who emphasized the fact that stability needs to be understood in a more contextual way (see for example Larson, 1997). Thus, whereas inherent stability, as a property of an event within a particular key, appears to exhibit a rather static representation of tonal pitch, contextual stability conveys a more *dynamic picture* of the organization of a piece, in the sense that it tackles relationships *between* pitch events within and between constituents, according to their structural roles. Stability accounts for a ranking of events according to their position within an abstract scale, while dependency shows how events enter into construction with each another. Portraying a picture of structural priority, dependency appears to be a more comprehensive structural property, in the sense that it only accounts for a ranking between pitch events in terms of their structural importance but also for relationships that describe which events elaborate which other events.

On the other hand, there is a debate in the field of music theory around the idea of what *is* structurally prevalent when tonal music hierarchy is investigated (see Chapters I and II for general considerations about this issue). Elucidation of the status of structural priority is important if the aim is that structural priority becomes a valid concept to provide coherent explanations about the way hierarchical relations in general and dependency relations in particular are understood. On the one hand, when priority is considered, the emphasis is placed upon the idea of linearity as the foundation of music hierarchy. Accordingly, priority is placed on the scale-degree capacity of pitches to relate to each other establishing linear connections that recursively compose-out a basic structural foundation until the musical surface is reached. On the other hand, explanations about the origin of hierarchy are located in the domain of harmony; here, tonal structure is understood as being the result of the operation of structural functions.

According to the first view, dependency is mainly linearly derived and originates in the application of techniques of linear ornamentation to a basic voice-leading structure. While in the second approach, a dependent event and its head are linked by descriptions of their functional roles and not informed exclusively by inherent pitch properties, horizontal features or syntactic categories of tonal pitch events. More specifically, prolongation - as a harmonic phenomenon - is associated with the harmonic identity of the events within a key. Therefore, it is the event's scale-step identity, *understood as the root of each harmonic chord*, that is prolonged. In other words, in the second approach linear procedures do not generate by themselves awareness of hierarchic differences between events, unless structurally more important events are somehow represented at a harmonic level.

Tacit analytical judgments of priority occur when, for example, a linear event that is located in the middle of a step-like moving line is labelled as a passing note. This leads to the question of the hierarchical status of the relationships between tonal events and the conclusion that when a linear analysis is performed, it is usually the case that analytical decisions about priority seem to be not just mere consequences of the linear status of the events in question but the result of the analysis of their structural roles, given that events have different meanings according to the extent they reflect different syntactic hierarchies. So, in the end, analytical decisions are about the structural roles events are *understood* to perform.

Anyway, as we noted before, Schenkerian tradition developed a linear conception of musical structure based on the unfolding of the tonic chord (see Chapter I for a detailed explanation of Schenkerian theory). Therefore, in Schenkerian tradition, it is acknowledged that the theorist's intuitions about structural priority are based in the first place on the idea that a tonal piece exhibits an event's organization in which subordinate events elaborate and thus prolong those superordinate events upon which they are dependent. However, it is frequently agreed that Schenker based the analysis of underlying structure mainly on linear aspects of musical compositions. And also, that hierarchy, to the extent that it is linearly derived, is mainly the result of the application of linear elaboration techniques to a basic voice-leading structure. Consequently, in Schenkerian theory, dominance-subordination relations between events and prolongational phenomena are linearly conceived. The linear elaborations substantiated through the voice-leading processes shape the hierarchical organization of tonal pieces and are assumed to be the basis of our understanding of such organization.

Even acknowledging the capacity of Schenkerian analysis as an agent to express our intuitions about the hierarchical organization of a musical piece, some theoretical approaches consider that the voice-leading patterns are not the sole source of our understanding of the underlying hierarchy. Accordingly, it is posited that the linear dimension of a piece's organization might exist in parallel with other pitch dimensions, for example the harmonic one (Horton, 2003).

However, as was highlighted in this thesis (see Chapter I), it does not follow from the Schenkerian conception that hierarchical relations are derived from a characterization of pitch events as entities that are *exclusively* constrained by the inherent stability provided by their syntactic scale-degree status, in order to be subjected to horizontalized composing-out procedures. Schenker's linearity is embedded and mutually informed by a contrapuntal-harmonic conception of music (see quotation in Chapter I, section II.2.2.1). The *Ursatz* is itself a contrapuntal-harmonic entity. The voice-leading arrangement of a tonal composition subrogates the attainment of structural goals to the rules of the fundamental progression. Schenker's descriptions of how events enter into construction with one another in order to generate linear progressions at the background, the middleground and the foreground levels are driven by contrapuntal-harmonic concerns.

An issue that emerges from the previous explanation is whether those linear processes that describe the way subordinate events depend on their leading events have the capacity to provide a consistent picture of our understanding of dependency relations as a prolongational phenomenon. Prolongation, according to Schenker, describes a structural property by which the musical influence of a structural event is extended beyond its specific sounding duration through the operation of linear processes (see Chapter I). As Horton (2003) emphasizes, the question of how a tonal structure is formed appears to be

different from the problem of what determinants are responsible of our understanding of tonal relations. In this case such differences should take into consideration, on the one hand, the factors that generate such dependency structures and, on the other, how the sense of dependency that gives rise to the understanding of some events as prolonging others is experienced, if it is experienced at all. According to the linear analysis, if one event is subordinated to another at a given level of the hierarchy, it means that the former elaborates the later; as this process is reproduced to some extent throughout the underlying hierarchy, an event that is a goal at one level may be an embellishment of another event at a higher level. It might be possible that the extent to which prolongation is cognitively experienced depends on our capacity to reassess the status of pitch events implicitly, according to the clues provided by the contrapuntal-harmonic underlying organization of the voice-leading unfolding throughout a piece of music.

Therefore, consideration of linearity should not appear to be just a question of scale-degree identity but also of the structural roles performed by the linear events at each level of the hierarchical organization of the musical piece. And on the basis of this feature, hypotheses about the way dependency - as a syntactical phenomenon - is experienced, might be elicited.

Bringing these ideas together, prolongational phenomena may account for a linear representation in which one event is subordinate to another because the first involves an embellishment operation that elaborates the second (such as a neighbour note or a passing note); or prolongation might inform harmonic-functional dependency representations in which one event is subordinate to another because the first functions harmonically with respect to the second. In spite of the theoretical disagreements about the epistemic status of hierarchical tonal organization, it seems that the mere *existence* of a prolongational

dimension in musical structure is taken for granted in the musicological and music-analytical literature. Overall, analytical descriptions acknowledge a distinction between the above-mentioned two separate dimensions of pitch organization (linear and harmonic) and also explicitly recognize their parallel co-existence.

If the aim is to understand which aspect of the organization of pitch events characterizes their syntactic status in the experience of linearly generated structures, it is necessary to create appropriate experimental conditions to test participants' sensitivity to those structural events. It is also necessary to inquire about those aspects of the identity of pitch events that characterize dependency relations. By exploring dependency relations, those linear templates revealed by the derivation of pitch events will be identified. At the same time, the harmonic dimension of such events and the capacity of some of them to prolong others will be acknowledged. It is assumed that the linear dimension operating in tandem with the functional dimension is what produces tonal coherence in music.

The cognitive experience of prolongational structure as a feature that obtains its hierarchical status by means of its dependent organization will be investigated in the context of experimental situations involving attentive listening to music. In what follows, the problem of attention in general and of to attending music in particular will be presented. At the end, a hypothesis will be posited of attending to the prolongational structure of a tonal piece as a syntactic feature.

IV.3 Constituency and attention to music

IV.3.1 The concept of attention in cognitive psychology

In Chapters II and III, the problem of attending to structural features of music was initially approached and one of the main music psychological frameworks that tackles this issue was reviewed (see Jones, 1979; Jones and Boltz, 1989; Boltz, 1989 and Boltz, 1993, theory of music attending in Chapter II). Attending to music from the point of view of its relationship to constituency in tonal music will be reconsidered indue lower. But let us first take a brief view of the general problem of attention in cognitive psychology.

Overall, the notion of attention is linked to the ability to focus or concentrate on aspects of reality. Inherent to the study of attention is our capacity to take into account the attributes of the object to which we are attending. In this sense, it is expected that attention will be paid to the relevant features, ignoring the irrelevant. However, a problem comes up when what is and what is not relevant is considered. This problem is related to the assessment of the enormous environmental variability within a given attentional situation.

An essential characteristic that emerges from the process of attention is therefore related to our capacity to balance our sensitivity to new incoming information in order to gain appropriate levels of concentration on what is relevant. Given the limited capacity of our brains to handle incoming perceptual input effectively as a whole, we appear to attend just to fragments of the input at a time. It seems that only a portion of the input directly affects our behaviour and is related to the events that we remember.

It seems that in the experience of organized sound, several memory processes have a functional propensity to allow us to process information in an increasingly complex way that initiates with echoic memory and early processing, and progresses afterwards into

more sophisticated forms of perceptual categorization (Edelman, 1992). By means of extraction of individual acoustical features and perceptual binding of those features into coherent auditory events (Bregman, 1990), our grouping mechanisms progressively organize events in terms of similarity and temporal proximity. Finally, conceptual categories, that result from interactions between bottom-up and top-down processing, are internalized. (Snyder, 2000; Mc Adams & Bigand, 1994); (see Chapter II for a description of the interaction of both systems in tonal processing).

The process of attention is regulated by two factors: i) retention in memory of potentially relevant information by short temporal spans, in order to reduce the possibility of losing valuable input and ii) filter of irrelevant information with the aim of avoiding overcharges in the memory system.

Another problem related to the way attention operates concerns the inherent unpredictability of the environment. Coping with it requires us to balance attention allocation in goal-oriented information while concentrating on the continuous stream of changing information coming from the environment.

Due to all that was noted above, attention is, more than a unitary concept, a cognitive dimension that relates to a variety of cognitive processes that utilize different brain structures (Peretz and Zatorre, 2004).

The functions of attention can be analysed from three distinct, although closely related perspectives: i) the scope of attention capacity, which refers to our limited attentional resources; ii) the relationship between attention and action control, which concerns the ability of attention to monitor and coordinate our behaviour and iii) the relationship between attention and information integration, that refers to the compromise of attention

when multiple sources and levels of perceptual information are integrated (Medin and Ross, 1997).

We possess internal mechanisms to keep a sensory register of the incoming information by means of which we automatically store a great amount of input signal during short temporal spans in order to process it. These mechanisms are called iconic (visual) and echoic (sonic) memories (Snyder, 2000). The limited capacity of short term memory and the commensurate amount of complex information to which we are continuously exposed, lead us to consider two ways of input processing. They involve, in turn, two distinct ways of managing our limited attentional resources: (i) serial processing, in which we process one element at a time and (ii) parallel processing, in which more than two elements are processed simultaneously. It seems that simple features may be abstracted and processed preattentively in parallel, without the need of selective attention; on the other hand, finer discriminations appear to require more focused processes, whether the processing is serial or in parallel. But, according to what was asserted in Chapters I and II, complex serial processing and behaviour in hierarchical organizations are by no means the result of simple occurrences of the type "A produces B"; on the contrary, complex serial behaviour is strongly dependent on integrated sets of cognitive structures previously planned and organized, within which sequences of occurrences are orchestrated. This is evident, for example, in the occurrence of speech errors that are the result of the subject's anticipation of words that come later in the sequence of the utterance.

But how do we manage to select the relevant information and at the same time to exclude the irrelevant from the focus of attention? Moreover, as far as music is concerned, what is relevant and what is irrelevant when we attend to it? It is a question to determine whether the human parser is serial, meaning that one structural interpretation is available at a time,

or whether it is parallel, maintaining more than one interpretation sometimes, according to particular circumstances (Gibson and Pearlmutter, 2000). Theories that hypothesize central processing systems postulate that selective mechanisms are responsible for successively filtering relevant features, specific properties and contextual characteristics of the input according to their relative priority, even though they are not purposefully the subject of attention (see Chapter II for general descriptions and applications to music modeling). Information assessed as of high priority receives further processing, leading to higher stages of meaning apprehension. Priorities at higher levels are continuously monitored by the attentional system, in response to a subject's expectations. Similarly to what happens in language processing, if the subject is expecting a specific continuation according to whether he/she has previously assessed some information as of high priority, the structural arrangement of music unfolding primes his attention to expect some developments more than others, according to the structure of musical idiom.

The way attention proceeds when checking for high-priority events is made possible thanks to the involvement of the long-term memory system. Transferring this idea to processing music hierarchy, if attention to musical structure is regulated by a hypothetical interaction between the underlying organization and details of the musical surface, then high priority decisions will be based in the way attention tackles information interacting between bottom-up and top-down processing, that is to say, by the way in which familiarity with the idiomatic syntax of music elicits checking decisions about structural priority, while music unfolds in time.

Multiple resource theories, an alternative to the theories of limited capacity, posit that despite the limited capacity to perform cognitive activities, our limited attentional resources can be distributed in multiple pools, allowing us an adequate amount of control

over the way our capacity is distributed across different tasks' components. Here attention will depend on the extent to which the task draws on the same and/or different resource pools (Medin and Ross, 1997).

Besides, it seems that learning plays a significant role in expanding attention and capacity limitations. Theories of automaticity (Logan, 1988; 1992) postulate that automatization is construed as the result of acquiring a domain-specific knowledge base, formed as separate representations of numerous instances of exposure to a given experience. Automatic processing relies increasingly on retrieval memory. Therefore, performance improves with practice associated with the accumulation of stored instances. According to this approach automatic processes are not completely involuntary but, as a consequence of practice, they become faster and require less time to perform the acts of control. One challenge to these theories is to specify what an experiential instance consists of and, as a consequence, to describe how generalization takes place.It seems that the more a process has been practised, the less attention it requires.

It is not just practice which reduces attentional demands, but the kind of practice involved. Specifically, whether practice yields automaticity depends on the *structure* of the task. In the case of music, for example, given that an episode of experience of a particular musical piece exhibits, to a certain extent, commonalities with other episodes (that is, with the experience of other musical pieces) but is also unique in some other respects, it is an interesting question for music psychology as to what fragments of experience memory and automaticity are founded. It is probable that in the case of a tonal musical composition those aspects related to the structural invariants of the tonal idiom become clues to prompt automaticity in the context of continuous exposure to the musical style and, as a consequence, orient attentional processes.

133

But if the interest is to study the problem of attention to underlying musical structure, the issue of the relative relevance of musical information still remains. As it was seen, theories of reduced memory capacity deal with the problem of attention in terms of the relationship between relevant - more salient - and irrelevant - less salient - information and their degree of automatization; according to them, attention to music would be the result of whether irrelevant information interferes with or has no effect on the processing of relevant information, depending on whether relevant information, irrelevant information or both have become automatized. Therefore, automaticity provides only a partial response. On the other hand, multiple resource theories, while positing different resource pools for different modalities (visual, auditory, etc) elicit the problem of the interference between different levels of a complex task in one modality, given the limited resource capacity of each different pool. Therefore, the problem of how the attending process is related to the capacity of the system to account for different levels in the hierarchy of music organization remains.

Selective attention is also necessary if the aim is to allocate the focus on relevant information. It involves focusing on a specific aspect of a *scene* while ignoring other aspects. It can be conscious, as when one deliberately chooses to attend, for example, to a particular aspect of a sonic scene (focusing on one voice during a social meeting in which many people are speaking) or unconscious; that is, when one thinks one has attended to the whole scene but only the processing of a partial area has occurred, the rest of the information having been taken in a general manner, as a "gist". Selective attention may select objects as well as or instead of locations; that is to say it may be object-centered and/or location-centered. Object and spatial dimensions are not necessarily mutually exclusive. In fact, it is possible that object-based attention is implemented by first

determining an object's location and then allocating attention to that location. These assumptions, relative to the visual domain, can be translated into the auditory domain, where it is necessary to consider the temporal dimension in the analysis of attention. Given the temporal nature of music, one can assume that focused attention results from the combination of object-centered (musical events) attention and serial order of appearance (events succession) attention (see in Chapter II relationships between the concept of attention with theories of musical processing).

But in order to locate attention in the context of the experience of underlying musical structure, selective attention cannot consist just of a split into two pools, that of relevant and that of irrelevant information. It is necessary to find a solution to the way relevance is assessed, taking into account that underlying structures take the form of one-to-many relationships. A solution that assigns an all-or-none psychological status to relevance, splitting information in terms of relevant versus irrelevant, is not useful to the purposes of this study. If the interest is to see if this kind of one-to-many relationship somehow orients the way musical structure is attended to, it is necessary to find an explanation that accounts for the complexity of this type of structural organization.

In sum, attention should be thought of as enabling people to organize relevant information and not solely as a limitation of their capacity for processing information.

In the concluding section of this chapter a hypothesis about the experience of attending to the underlying structure as a constituent organization will be presented.

IV.3.2 Attending to music hierarchy as a constituent organization

Music is a temporal art. As explained so far in this thesis, this peculiar characteristic of music unfolding and the musicological and psychological implications it epitomizes, have

been considered extensively. A hypothesis concerning the experience of music as an underlying hierarchy could be developed, positing that the underlying hierarchy could be understood in terms of a constituent organization, in which prolongational features govern its unfolding. In order to do this, it is necessary to consider the way attention interacts with music organization in the context of its temporal deployment. In Chapter II, some of the main static and dynamic models of music cognition, which analyze this problem from different perspectives, were reviewed. In particular, one of these theoretical frameworks approaches the problem of music attending (see Jones's theory of music attending in Chapter II).

As was seen in the previous section, attention seems to facilitate cognitive processing, helping organize relevant information. At the same time, it seems that the problem of attention is influenced by event structure, that is to say, by the nature of the stimulus events that constitute the temporal interval (Brown and Boltz, 2002). The interest here is to apply the concept of attention to a hypothesis of the experience of the prolongational structure of tonal piece's constituent organizations. If the underlying organization somehow orients attention, interacting with the way cognitive resources are used to process the piece while it unfolds in time, it is expected that information processing varies along the constituent unit according to its characteristic internal organization, ruled by dependency constraints.

In other words, the subjects' sensitivity to the components of a constituent unit will account for their relative relevance. A structural constituent has been assumed as a complex organization in which one event (the prolonged head) is extended throughout the constituent unit throughout the prolonging events. An interaction between the particular arrangement of the stimulus and the ways it is experienced will involve orienting attention allocation and priming expectations about the relative importance of the constituent events.

136

As a consequence of such interaction, decisions about the structural priority of events will be elicited. In so doing, attention will i) split into more relevant (the constituent's heads) and less relevant (the other tones of the constituents) events, priming expectations about the capacity of less relevant events to extend the priority of the more relevant constituent's head, and ii) map those aspects of music unfolding of the constituent unit onto appropriate control parameters for action. For example, information relative to the structural priority of the constituent events within the underlying structure can be used to monitor the attentional effort, while processing of the unfolding constituent goes on. If, for example, time reaction responses were required from listeners that are attending to music, it would be expected that variations in their response latencies reflect the listeners' sensitivity to the relative structural priority of the focal points in which attention is allocated.

The prolongational status of the constituent unit might orient conscious and/or unconscious attention to the head of the constituent, and prompt a general 'gist-manner' integration (see above) of the rest of the information to the primary event (the constituent head), until a new constituent head appears that requires focusing attention allocation again.

It seems that interaction between the ways music unfolds and our attentional resources allows anticipation of the future course of events, including the moment in which a given temporal interval is expected to finish. Assuming that the surface structure of a musical piece represents a dynamic organization, a hypothesis of music experience of the underlying hierarchy in terms of an interaction between its structural compositionality and the way attention is performed can be derived. It will posit that attention will vary according to the structural importance of the musical event that is being focused upon, and that decisions about what is actually focused upon will be related to the constituent nature of the prolongational structure of the piece. If attention is modelled by prolongational

constituency, it is expected that dominant, prolonged events become cognitive reference points to which the focus of attention is allocated, and subordinate, prolonging events, that are associated to the former, orient expectations of continuation and/or closure, according to the linear constraints of the underlying voice-leading.

According to Jones & Boltz (1989), highly coherent events support a kind of future oriented attention. If the underlying voice-leading is a structural framework that provides a high degree of predictability it will orient attention in such a way as to generate expectations of how and when the constituent units are going to finish. If the notes of the underlying voice-leading are temporally organized in coherent ways, thanks to future oriented attention (see detailed explanation of this concept in Chapter I), listeners will frequently anticipate not just which notes are about to come but also the temporal moment they are going to occur.

Hierarchical structures exhibit high temporal coherence. For example, it has been found that listeners are inclined to attend in the first place to higher levels of the formal structure in experimental tasks that require attention to form in musical pieces (Stoffer, 1985); such an attentional mode manifests itself by means of quicker temporal responses (see details of studies that measure subject's reaction time in the following section of this chapter). Thus, the most prominent temporal levels become markers for the most salient non-temporal information. If dynamic attending is organized around those features it will select where to allocate the focus at the moment-to-moment representation of the ongoing musical structure. Consequently it is expected that subjects' responses in different locations of the constituent hierarchy will reflect such differences.

Events' organizations with low temporal predictability prevent people from anticipating their future development, forcing them to attend locally to adjacent elements with the

intention of organizing such non-structured information. On the other hand, future oriented attending might support a hypothesis of constituency as a one-to-many relationship with a high level of predictability. In this case, fulfilment of temporal goals would derive from listener's dynamic interactions with patternings of the constituent structure of the piece. It is assumed that the most important events become structural markings that guide attentional sensitivities to the satisfaction of expectancies throughout time.

Two features contribute to the definition of hierarchical temporal structures: the temporal markers and their distribution. Beginnings and endings are intrinsically marked by structural changes. The way events fill time intervals at different hierarchical levels and the way a person attends to them can affect time estimations. Studies that required subjects to compare the total durations of two melodies within pairs of stimuli, found that duration judgments differed significantly as a function of the kind of information filling the temporal periods, in spite of the fact that no durational differences really existed. Participants judged patterns that seemed to end too soon to be short and those that seemed to end late to be long. Attending to music over time involves people relying directly on event's structure and adopting alternative temporal perspectives in order to fulfill different structural goals. Listeners seem to be relatively accurate in predicting endings of hierarchical melodies as a result of the interaction between their expectations - arising from future-oriented attending - and the temporal coherence of events' structure, (Jones and Boltz, 1989).

Given the hierarchical nature of underlying structures, it should be possible, accordingly, to determine functional markers at different levels of the hierarchy. In theories of dynamic attending, markers are defined relationally as contextual change. Non-temporal information, such as the relationship between two pitch events in a linear voice-leading

arrangement, has some potential to establish a significant time interval, whose boundaries can be experienced as cognitive reference points. The underlying hierarchy of a musical piece could be assumed as a structural organization in which distribution of markers reveals nested levels that are consistently related to each other by certain types of arrangement. For example, the location of the constituent's head and the end of its syntactic extension might delimit constituent boundaries which cognitive status deserves investigation.

IV.3.3 Assessment of attention to the syntactic structure: the click technique

A series of seminal experiments in linguistics used the click-detection and click-location paradigms to study linguistic information processing. The click technique consists of superimposing an extraneous signal on the linguistic stream of events at different focal points of its constituent structure. In click-detection experiments participants are required to detect the click producing a simple motor task, for example, pressing a key such that their reaction time (RT) to different click positions is assessed. In click-location studies, it was expected that participants showed sensitivity to constituent boundaries migrating the position of the clicks to such locations. Both paradigms have also been used in musical experiments (Gregory, 1978; Sloboda and Gregory, 1980; Stoffer, 1985; Berent and Perfetti, 1993).

Fodor and Bever (1965) ran experiments with spoken sentences using the click-location technique. Results indicated that there seems to be a tendency to represent events that are highly related as uninterrupted, in such a way that an extraneous signal (the click) is pushed to a point where structural links are weak. A click migration effect was thus produced. They concluded that the constituent structure of spoken sentences is represented

psychologically in terms of relations of structural proximity. The immediate constituent functions as a perceptual unit in the processing of sentences. Additionally, the results of this study demonstrated that clicks are more accurately located at major constituent boundaries than in the interior of the constituent unit underlying a hypothesis that processing load is greater during the constituent unit than at the end of it.

Holmes and Forster (1970) used the click detection technique with spoken sentences to test Fodor and Bever's assumption of differences in processing load. They posited a hypothesis of heavier processing load during a major constituent than at the end of it, and a concomitant slower detection time for clicks located inside and faster detection time at the boundary of the constituent unit, respectively. Subjects listened to spoken sentences and they had to press a key as soon as they detected the clicks. Results confirmed the hypothesis of processing load, given that RTs were faster for clicks located at the major syntactic boundaries compared to RTs to clicks located within the sentence. Also RTs were slower in the first part of the sentence than in the second one, indicating that RT is a function of the amount of interference from sentence processing. That is to say, at the beginning of the sentence, unpredictability about the sentence's future deployment is maximal, while predictability increases by the middle to the end. This of course is congruent with the above-mentioned differences in decoding load.

Positing similarities between the perception of music and speech, studies on music (Gregory, 1978) used the click location technique by involving subjects in the task of listening to short sound sequences of six notes over which clicks had been superimposed at particular locations within their grouping structure. Instructions invited subjects to look at their written representation in a musical staff and to point out the position of the click. Results found some differences between speech and music processing: in speech, clicks

141

tended to be perceived early, but in music late responses were much more frequent than early ones. But more importantly, clicks showed a tendency to migrate to the grouping boundaries, suggesting that the perception of both speech and music seems to be organized in phrases.

Along with the investigation of parallels between language and music, Sloboda and Gregory (1980) ran a new experiment on click location by increasing the complexity of musical stimuli and presenting subjects two-phrase melodies of 16 notes. They placed several clicks before and after the phrase boundaries in order to investigate the participant's location of the signals. They treated the melodies in such a way that they were able to provide cues to test the effects of structural (harmonic) and physical (duration) markers in the participants' location of clicks. Subjects were required to indicate exactly when they heard the click by reporting the note that they thought was sounding simultaneously with the click. The results seemed to indicate that clicks were in general located after their real occurrence, except for clicks that immediately followed the phrase boundaries that were migrated backwards. In general, they demonstrated the click migration effect; and they also provided evidence that clicks seem to migrate perceptually toward the boundaries between structural groupings, and that physical and structural phrase markers affect click migration differently. Structural markers (in the harmonic sense) seem to affect migration when clicks occur after the phrase boundary, indicating that perhaps participants are unable to anticipate the phrase boundary on the basis of those cues alone. On the other hand, physical markers (longer notes) seem to exert a stronger effect eliciting anticipatory migrations. In summary, they provided evidence of the psychological reality of musical fragments, defined by music theory as phrase-structures, and gave an insight into the similarities of music and language processing.

In an extended replication and generalization of the results obtained by Gregory (1978) and Sloboda and Gregory (1980), Stoffer (1985) tested listeners' sensitivity to the formal structure of pitch sequences using click localization (experiment 1) and click detection (experiment 2) techniques. In experiment 1, materials were extended from previous experiments with respect to the number and order of phrase-structure boundaries which were tested. Artificial children's songs were used and 17 click positions and five-order boundaries were tested; boundaries were marked by changes in implied harmony, in melodic contour and also rhythmically by the addition of accents. The aim was to show a distribution of click localizations in which (i) the probability of locating a click at a given position is proportional to the order of phrase-structure at that given position, that is, the higher the structural order of the phrase boundary the higher the probability of locating the click, and (ii) the quantity and direction of the click migration is a function of the structural order of the phrase boundary where the click really appears. Similarly to what happened in speech processing, it was expected that migration would tend to minimize the number of processing units that the click interrupts. The higher the number of superimposed units at the click position, the higher the listener's tendency to migrate the click to the common boundary of those formal units. Stoffer obtained significant differences between different click positions and demonstrated that the click's localization depends on the hierarchy of the structural phrase boundary where the click is placed. Results showed a tendency to migrate the click position to the main phrase boundary.

In summary, the results of experiment 1 indicated that locating a click in a melody is the consequence of a process in which the representation of the unit in which the click is placed is progressively reduced - through a top-down procedure - along the representational hierarchy. As low-level representations are less stable, it takes more time

to access them. Such instability in memory reduces the effectiveness in click localization at lower levels, while localizations at higher levels are more accurate and quicker. The last is true given that localization always takes as reference points the higher order boundaries that are more stable in perception. The deeper (in a top-down sense) the click is embedded in the hierarchy, the longer the time required to process it. The lowest level of representation is the level that is least stable in memory because it takes more time to access in comparison to the highest level of representation. This is the reason why localization frequencies and migrations are higher, the higher the structural order that acts as reference point. This conclusion is interesting in the sense that it accounts for the cognitive stability of syntactic structure in music (see Chapter II for a description of stability conditions in music cognition).

In experiment 2, on click detection, the aim was to test if the binary partition of a structural unit is a kind of prototype that is used by non musicians to segment complex phrases regardless of their actual structure. Thus, subjects who were not familiar with phrase complexity would use a binary division heuristic as a prototype in the resolution of the task. Therefore, musicians and non musicians were tested to compare their behaviour. Two phrase-structures that are characteristic of children folk songs were used. One of them consisted of three phrases of four measures each, in which the first eight measures formed the prototypical binary phrase structure. The second phrase-structure was a derivation from the first one containing the same number of measures but a different internal phrase organization. Both phrase structures were arranged so that they presented noticeable differences between them in the order of phrase boundaries at the same position.

Overall, there were expected differences in RTs between clicks located at those positions, according to the listener's representation of the differences between the phrase-structures.

144

It was predicted that the higher the structural importance of the boundary at which the click is located, the shorter the response latency to the click. Differences in latencies at a position determined according to hierarchical order were expected in musicians. Such differences were not expected in non-musicians at the beginning, because it was hypothesized that non-musicians would use a binary division heuristic; however, it was expected that at the end, differences between groups would be reduced, due to learning by non-musicians during the experimental sessions how to segment a non binary phrase structure. In general, the results showed significant differences between click positions depending on the phrase structure, independently of differences in musical competence between the two groups of participants. In particular, a significant difference in latency was found between the click position at the main boundary (faster RT) and the rest of the click positions in both phrase structures. The results also provided some evidence of the correspondence between response latency and structural order of the boundary at which the click was located: differences in latency were found in some click positions but not in all of them. Analysis of results, splitting responses between the two groups of musical competence, were in agreement with the prediction that processing of the binary structure showed no difference between the two groups during the experimental sessions; in the case of the non-structure, however, musicians showed the opposite behaviour to non-musicians at the first session, but there were no differences in the last session between the two groups. Non-musicians changed their segmentation behaviour during sessions, showing evidence of the acquisition of musical knowledge.

Berent and Perfetti (1993) also used the click detection technique in an experiment that tested the on-line parsing of harmonic sequences. They were interested in investigating listeners' sensitivity to the process by which the representation of a musical piece is

constructed on-line (that is, in the moment-to-moment of a real time listening situation); in particular, they wanted to study listeners' decisions when they process non-modulating and modulating harmonic sequences. Using a divided attention paradigm with music listening as the primary task and click detection as the secondary task, they hypothesized that increasing the complexity of the primary task would result in an increase in response latencies to a click presented immediately after a modulatory shift. They posited a parser problem, defined as the limitation of the listener to arrive at a unique immediate representation decision when the musical input exhibits ambiguity; as a consequence of this problem the parser has to make a decision when the musical grammar is still unable to suggest a solution. If premature decisions - that are made as long as the music unfolds - have to be reassessed, the listener will experience, according to Berent and Perfetti, a parsing failure that requires reinterpretation to be made. Reinterpretation suggests that a single musical event may have different representations at distinct time points throughout the listening process. In the experiment it was assumed that reinterpretations demand more attentional resources with the concurrent increase of latency responses. The method used in the experiment assumed a limited attention pool in which if demand increases performance decreases (see section IV.2.1 in this chapter for a description of this concept). Therefore, click detection as a secondary task would decrease in accuracy as a result of increments in the load created by complexity of the listening primary task.

Using modulating and non-modulating harmonic sequences (the latter as control for the former) they tested the ability of click detection to reflect differences in the cognitive load required to process the sequences. It was predicted that in the modulating region responses would be slower than in the non modulating one. Their results answered the assumption of the direct relationship between increments in cognitive load and musical complexity. Also,

it was demonstrated that increments in cognitive load can be detected in the performance of a secondary task by means of the production of longer response latencies. Finally, the results supported the validity of the click-detection paradigm as an on-line method. The paradigm seems to be sensitive both to the global structural complexity and to detection of changes immediately as they occur.

Berent and Perfetti also found expertise-based differences between musicians and non-musicians in their attention-allocation strategies while processing the samples. It seems that non-musicians treated monitoring of clicks rather than listening to music as the primary task, allocating their attention mainly to click detection.

In summary, the use of the click technique both in linguistic and in musical experiments has proved to be a useful strategy to test participants' processing of constituent phrase-structures. There is a strong tendency either to react faster to (click detection) or to migrate (click location) those signals located near higher order boundaries. Constituent structures of spoken sentences and of musical phrases are represented psychologically in terms of relations of structural proximity. That is to say, there seems to be a tendency to represent events that are highly related as uninterrupted, in such a way that the extraneous signal (the click) is pushed to a point where structural links are weak. Besides, click location accuracy is higher at major constituent boundaries than in the interior of the constituent unit, underlying hypotheses about variations in the amount of processing load during the constituent unit. Results also indicate that the processing of constituency seems to progressively reduce the representational hierarchy in a top-down way, because localizations at higher levels are the ones that are more accurate and quicker. Finally, the click technique has also been used with success as an on-line method to test the experience of harmonic constituency; results demonstrated that the production of longer response

latencies accounts for increments in the cognitive load when click detection is used as a secondary task.

However, there are several shortcomings that emerge from the use of this technique. In the first place, click migration behaviour emerging from click location studies may not be the result of factors that operate during the processing of the constituent structure but instead could originate in events that occur after the perceptual analysis has taken place. This is due to the fact that the participant is required to listen to the stimulus first and to produce the response afterwards, simultaneously operating with written representations of sentences and musical fragments. This delay in the subject's performance compromises memory. In order to overcome this shortcoming, some experiments tried to minimize the effect of memory by involving subjects in tasks that did not require the use of written representations. They were replaced by auditory input alone and the production of a verbal response as close as possible to the hearing of the stimulus. But in doing this, other potential sources of interference were included; they emerged from the task required of participants: they were involved in running comparisons between pairs of melodies and pairs of click positions (Fodor, Bever and Garret, 1968, in Holmes and Forster, 1970). Therefore, it seemed to be intrinsically impossible to separate the memory effects from the perceptual process *per se* using the click location methodology. But the most important shortcoming of click migration performance emerges from the experiments run by Holmer and Forster. They demonstrated that the hypothesis of Fodor and Bever about the nature of migration was more a consequence of the lack of participants' security in their judgments than a derivation of their representation of the constituent boundaries. Uncertainty in perception and concurrent migration would emerge from the amount of processing within the constituent.

In the case of musical studies, the use of staff notation also implies potential sources of interference such as, for example, the degree of ability to use reading strategies and previous musical knowledge. Memory represents another source of interference. Control over the last factor was attempted, for example, by presenting subjects with very short fragments of sound sequences (Gregory, 1978) with the concurrent loss of musical meaning. That is to say, if music has meaning at all, there is a strong difference between the syntactic meaning conveyed by a speech sentence and the syntactic meaning emerging from a six note pattern. Although the use of more extended sound sequences that formed musical phrases (Sloboda and Gregory, 1980) may be considered an improvement of the previous constraint, experimental stimulus presentation involved sound sequences performed by a sound machine almost always with equal inter-onset interval and in a synthesized timbre. The problem with these impoverished materials is that they do not represent ecologically valid musical examples from which significant conclusions about the musical experience of the syntactic structure can be derived.

Finally, click studies were used to test *phrase-structures and harmonic structures* as constituent organizations using musical samples with a dubious musical value. In our studies hypotheses concerning *prolongational structure* will be posited using real musical fragments.

IV.4 Summary: the assumption of prolongation as a constituent in music attending

In this chapter a hypothesis about the syntactic nature of prolongation in tonal music was posited. It was also proposed that this assumed syntactic component of prolongation might

to a certain extent drive the listener's sensitivity to the underlying structure during music attending.

A syntactic analysis of tonal structure takes into consideration the operation of two types of systems: constituency and dependency. Constituency is the feature of syntactic organizations in which structural units are formed by events that group together on the basis of sharing a structural characteristic. Constituent units combine with each other to form concatenated streams of events. The process of parsing a musical surface into constituent units is closely related to the identification of the constituents' *boundaries*. Dependency relations, on the other hand, account for the fact that some elements are subordinated to other superordinate elements on the basis of their structural qualities. Structurally speaking, dependency is therefore that aspect of an object's organization that exhibits a hierarchy of relations between leading and subordinate components throughout the whole network of events. This hierarchy of relations between events is manifested in a constituent unit by means of the relationship between the head (the most important event in the constituent unit) and the subordinate elements of that unit. In constituency systems, then, a unit is divided into other units; in dependency systems a unit governs other units. Both systems are strongly interdependent in syntactic organizations.

The hierarchical structure of music exhibits two features that emerge from constituency and dependency relations: precedence and dominance. Precedence deals with the temporal nature of music, requiring that some events appear before than others; precedence displays order in the concatenation of successive constituents and defines the linear structure of the piece. Dominance, on the other hand, concerns the hierarchical grouping of the constituent structure; in order for discrete events to be integrated into more comprehensive syntactical units, some of them must govern the operation of others.

To hypothesise prolongation as a constituent organization is to understand the relationships between events within a constituent unit in terms of a structural priority in which the head is prolonged by the dependents in such a way the constituent as a whole projects the structural quality of its head.

In this thesis a hypothesis of the experience of a constituent component of the prolongational structure will be investigated in an experimental situation which involves attention to music. Overall, the notion of attention is linked to the ability to focus or concentrate on aspects of reality. Given the limited capacity of the attentional system, a crucial problem of attention is related to the estimation of what is and what is not relevant in the complexity and variability of the information to be processed in a given situation. Attention enables people to *organize* complex information. In the context of the experience of prolongational structure it is assumed that prolongation takes the form of one-to-many relationships. If attention is modelled by prolongational constituency, it is expected that dominant, prolonged events become cognitive reference points to which the focus of attention is allocated, and subordinate, prolonging events, that are associated with the former, orient expectations of continuation and/or closure, according to the linear constraints of the underlying voice-leading. Beginnings and endings in constituent organizations are intrinsically marked by structural changes. Thus, the location of the constituent's head and the end of its syntactic extension will delimit constituent boundaries whose cognitive status deserves investigation.

The click technique was used with success in linguistic and in musical experiments to test participants' processing of constituent phrase-structures. There is a strong tendency either to react faster to (click detection) or to migrate (click location) those signals located near

higher order boundaries. In the following chapter, experiments that tested music attending to prolongational structure using this technique will be reported.

CHAPTER V

EXPERIMENTAL STUDIES PART 1: PROLONGATION AS A LINGUISTIC CONSTITUENT

Introduction

In this chapter, several investigations concerning the psychological reality of the prolongational structure in tonal music are presented. The cognitive status of structural boundaries is explored in different types of constituent organizations, in order to look for experimental evidence of their psychological reality during music attending tasks.

Previous research on prolongational structure investigated its cognitive status (i) involving participants in the use of similarity judgments to study the perceptual categorization of prolongations in tonal melodies (Martínez & Shifres 1999a, b; 2000; Shifres & Martínez 1999, 2000a) and (ii) requiring subjects to estimate the degree of correspondence between a melody and its rendered reduction to study the effects of the performer's actions on the listeners' representation of prolongation (Shifres & Martínez 2000b). A hypothesis of perceptual rivalry between the attributes of the musical surface and those of the underlying voice-leading was tested, indicating that prolongation is an attribute that is experienced (Martinez & Shifres, 1999c, and see Chapter III for a detailed description of these experiments).

From this conclusion, we can formulate the following general research assertion: that prolongational structure has cognitive status because it is used in similarity judgments and estimation of goodness of fit tasks.

Although the pertinence of those experimental paradigms to the study of prolongational structure was demonstrated, it was also emphasized that listeners' responses could be influenced by conditions of perceptual asymmetry and changes in the context, and that the particular cognitive task in which the participant was involved could have an effect (Martínez, 2000 a, b, 2001).

The aim of the studies presented here is to go further in exploring the psychological plausibility of prolongational structure, investigating its cognitive status in other cognitive processes involved in the experience of tonal music.

The experiments that will be reported seek behavioural evidence of listeners' sensitivity to prolongational structures, posing a hypothesis of constituency to the underlying voice-leading, and analyzing the cognitive status of the pitch events' organization in musical pieces under experimental conditions that combine *analytical and future oriented attending* (Jones and Boltz, 1989).

To assign a constituent function to prolongation (see Chapter IV for a description of the concept of constituent function) is to assume that the voice-leading unfolding of a musical piece can be subjected to a parsing process in which the linear continuity is represented in terms of syntactical units with beginnings and endings. Therefore, posing a constituent status to prolongational structures is to find in the voice-leading unfolding *boundary locations*.

Overall, the concept of constituency is intrinsically related to the concepts of relative closure and boundary segmentation. Discovering words embedded in a mostly continuous speech stream is one of the infant's first tasks in language acquisition, which is related to the capacity to process grammatical features of speech in order to identify word boundaries

(Saffran, 1996). Similarly, grammatical features of tonal music, derived from the analysis of embedded musical organizations, inform the constituency of music. In formal phrase-structures, for example, structural markers such as harmonic progressions, long duration notes, and metrical emphases or silences, among others, contribute to establish formal boundaries and to convey a sense of relative ending to constituent units (see more details about constituency in Chapter IV). There are cases in which segmentations between units are not so clear, such as in elided structures in which the last note of a formal unit and the first note of the following unit are the same. Therefore, the psychological reality of boundaries in constituent organizations is dependent on the relative degrees of closure and the concurrent segmentation between units in the internal configuration of musical pieces.

It is an assumption of constituent theory that the embedded and linear features of tonal organizations rule the relationships of structural components of a musical composition, providing listeners with different cues that allow them to experience sensitivity to musical closure. Experimental work in music cognition has found developmental and behavioural evidence of this cognitive process (Serafine, 1988; Stoffer, 1985), but contrasting results have been reported when closure was tested in different structural contexts and using different experimental paradigms. However, some of the experimental results that dismiss the value of long term tonal closure as experienced in music cognition (Cook, 1987) have been subjected to intense criticism (Gjerdingen, 2001).

On the other hand, the embedded organization that prolongational structure exhibits as a result of the voice-leading unfolding might be understood in a rather different way than the way constituent configurations are (see Horton, 2003). Posing a hypothesis of constituency to the prolongational structure means, nevertheless, that prolongation *must* exhibit, to a certain extent, a capacity to convey relative degrees of closure in the unfolding of the

prolongational units and that this capacity could be a clue to identify prolongational boundaries throughout a particular voice-leading arrangement.

If relative degrees of closure are experienced at different hierarchical levels in constituent organizations, it is an assumption of this thesis that they will be *also* experienced in prolongational structures.

Therefore, in this chapter four experiments will be presented. Experiment V.1 attempts to disambiguate the contrasting results found when sensitivity to closure in constituent structures was investigated. On-line sensitivity to closure in embedded and linear constituent structures is investigated using an on-line parsing paradigm, in order to find precise evidence that constituent boundaries *are effectively* experienced in terms of relative degrees of closure. Experiments V.2, V 3 and V.4 pose a hypothesis of constituency regarding prolongational structures and test the cognitive status of constituent boundaries using two experimental techniques: click-detection and click-location.

V.1 Experiment V.1: Sensitivity to closure in embedded and linear constituent contexts

V.1.1 Introduction

The concept of musical closure acknowledges a long tradition in the fields of music theory and psychology of music. Experience of closure is characterized as the sense of relative endedness achieved after listening to a musical fragment. The operation of some cognitive processes such as for example grouping, segmentation and abstraction of hierarchical levels, is informed by the relative degrees of closure that listeners experience in music (Lerdahl & Jackendoff 1983; Serafine 1988, Deliège 1987).

Embedded and linear features of constituent organizations provide the syntactical context in which relationships of structural components of the musical piece unfold. Guided by an internal, unconscious, grouping capacity, listeners spontaneuously segment the intrinsic organization of the musical input and form musical groups that are mentally represented as constituent-structure representations (Horton, 2003). The main characteristic of embedded structures is that their constituent units are heard in a hierarchical fashion (Lerdahl & Jackendoff, 1983, p. 13). That is to say, one unit subsumes or contains other units, being the latter dependent on and subordinated to the former. Given that this process of subordination/domination is assumed to continue indefinitely, it can be the case that one unit is more embedded than other, depending on the size of the level that dominates the constituent hierarchy in which it is included (Lerdahl & Jackendoff, 1983; Jones & Boltz, 1989). On the other hand, there are constituent constructions formed by the repetition of units of equal structural importance; in this case, the units are combined by means of an iterative process whereby a constituent is replicated several times consecutively. This mode of syntactic combination is called coordination (Horton, 2003). The extent to which whether a certain instance of repetition is understood as a coordination - of equivalent structural value and often tolerant to certain internal variation - or it alters the structure in which it is included is a question that deserves investigation.

The syntactical context of a constituent organization exhibits a potential capacity to bring listeners different cues that help them experience the characteristic sense of musical closure. Some of these cues arise from the degrees of continuation and/or finality of musical phrases. Grouping and closure interact providing continuity to the organized motion of musical pieces and making the succesive phrases related at higher levels of the

constituent hierarchy (Snyder, 2000). Therefore, closure is one of the main structural functions that emerges from the syntactic organization of musical pieces.

Given its prototypical character as a structural feature of tonal music, closure has been subjected to experimental research in the field of developmental and cognitive psychology of music. It was found that this cognitive process is present in music cognition at 7 years and is completely stable at the age of 10-11 years (Serafine, 1988). This deferred stability evidences not only a developmental trend but also the influence of an individual's enculturation, which involves awareness of peculiarities of the musical idiom whith which he or she is familiar. Given previously found contrasting results when closure was investigated, the current experiment tries to provide evidence of its cognitive experience in listening contexts.

V.1.1.1 Aim

The main objective of this study is to disambiguate above mentioned results that emerged from previous studies of tonal closure, in order to seek evidence of its cognitive status at boundary locations in constituent organizations. If this objective is fulfilled, that is to say, if it is found that *constituent* boundaries have psychological status, then it will be possible to pose a hypothesis of constituency to prolongational structures and investigate it looking for listener's sensitivity to *prolongational* boundaries. Finally, if evidence of sensitivity to prolongational boundaries is obtained, then, evidence for a constituent ontology of prolongation will also be obtained.

In the present study, listeners' sensitivity to relative closure is investigated in a real time listening situation in which the grammatical content of the constituent organization of a musical piece is analysed. Repeated motives that compose two types of constituent

organizations are tested: embedded structures and linear structures. Previous studies (Serafine, 1988; Stoffer, 1985) were run using either isolated excerpts of music, or deadpan tone sequences. Consequently, the use of these experimental devices avoided the study of an implicit grammatical component of the concept of closure. In contrast, the current study runs comparisons between the perception of closure in small, isolated musical units, and of those same units combined in a larger musical context. Besides, the interest is to explore closure in two different constituent organizations (embedded and linear) in order to know more about its psychological status.

V.1.2 Method

V.1.2.1 Subjects

188 undergraduate students from courses of music at the Faculty of Fine Arts of the Universidad Nacional de La Plata, La Plata Music Conservatoire and Berisso School of Arts volunteered to participate in the study. They were divided into two groups according to their musical experience: Group 1 - 101 students from undergraduate initial courses (28 were female and 73 male; average age 20 years) - and Group 2 - 87 students from undergraduate advanced courses (26 were female and 75 male (average age 22 years). Undergraduate initial students had an average of two years of musical tuition including ear training and instrumental lessons and advanced students had an average of five years of musical instruction with at least three courses on music theory. Members of both groups were randomly assigned to conditions 1, 2 and 3 of the experiment.

V.1.2.2 Stimuli

The first section [mm 1 to 36] of Chopin's Ballade No. 1, in G Minor, op. 23 was used to provide stimuli. The beginning of this beautiful piece presents a very interesting case of

embedded grammatical organization (see Figure V.1). Immediately after the introduction, a typical period phrase is unfolded, where it is possible to identify an antecedent (mm 9.4 to 13.3 [units a1-a2]) followed by a consequent (mm 13.4 to 17.3 [unit a.3]). In turn, the antecedent is organized in two minimal units (mm 9.4 to 11.3 [unit a.1]; and 11.4 to 13.3 [unit a.2], respectively). Therefore, a clear multilevel arrangement is created, which conforms to an *embedded hierarchical context* (Jones & Boltz, 1989; Lerdahl & Jackendoff, 1983). It is assumed that this grammatical organization is heard in a hierarchical fashion, meaning that the discrete elements that compose it are related in such a way that one element subsumes or contains other elements. In other words, the element that is subsumed or embedded is said to be subordinate to the element that contains it, being the latter the dominant or superordinate, and the former the subordinate. Interestingly, each of the units of the multilevel arrangement of the Ballade's first phrase period ends with an identical melodic motive (C-D-F sharp-B flat-A-G) arriving at the tonic G. At the beginning of the second phrase period there are two other units with this melodic motive (mm 17.4 to 19.3 [unit a.4]; and mm 19.4 to 21.3 [unit a.5], respectively). The last unit presents the characteristic melodic motive with a different harmony: a major tonic chord as a secondary dominant.

During the second phrase period it is possible to identify a new type of organization. This time, the concatenated elements compose an arrangement that is not hierarchic in the restricted sense of the previous embedded structure; instead, repetition (real or sequenced) conforms best to the so-called associational structure (Lerdahl & Jackendoff, 1983). In it, only one level is active, giving rise to a *linear context* (Snyder, 2000; Horton, 2003). The first pair of repeated fragments pictures a fairly literal repetition (mm 21.4 to 23.3 [unit b.1]; and mm 23.4 to 25.3 [unit b.2] although texture is slightly different in the second unit.

The second pair of repeated fragments presents a sequenced repetition (mm 27-28 [unit c.1]; and 29-30 [unit c.2]. In these linear contexts, the more salient grammatical feature is the locus of the units in the temporal sequence; it sets their function, and therefore, it is assumed that it contributes to the perception of their relative closure.

The whole excerpt was divided into fourteen fragments that formed syntactically meaningful units. Only nine out of the fourteen fragments (units a1-a5; b1-b2 and c1-c2, see Figure V.1) were considered for analysis of their relative closure. Units were assigned their syntactical status according to whether they were: i) *repeated* units in a hierarchically embedded context [a.1 to a.5] (note that, in fact, only the end of the unit is strictly repeated); ii) *repeated* units in a linear context [b.1 and b.2] and iii) *sequenced* units in a linear context [c.1 and c.2]. The rest of the units were presented to participants (i) to provide the whole musical context of the Ballade section and (ii) to produce unordered combinations with the other units in experimental condition 3 (see below).

These three types were investigated in three different experimental conditions:

Condition 1: *Successive units arranged in 'grammatical' concatenation*, consisting of the presentation of the complete original fragment with each of its fourteen units separated by a silent span of 2.5 sec. The inclusion of the silent lapse provided a clear indication of the piece segmentation without nevertheless interrupting the on-line listening of the musical flow.

Condition 2: *Successive units arranged in 'grammatical' isolation.* In this condition subjects listened to the complete fragment but its fourteen units were separated by a silent lapse of 10 sec each. This lapse, which is longer than the psychological present (Fraisse, 1974) was intended to separate the units with a temporal lapse long enough so as to break

Figure V.1 Score of the first section of Chopin's Ballade opus 23 no.1. The fragment was divided into fourteen units that are marked in brackets. Tested units are labeled in red.

the on-line sequential grouping that characterizes musical parsing during a typical listening situation without violating, however, the grammatical order of the units' presentation in the musical piece.

Condition 3: *Successive units arranged in 'ungrammatical' concatenation*. In this condition the fourteen units were separated by a silent lapse of 2.5 sec each, as in Condition 1, but this time the units were concatenated in another order than the original order of Chopin's Ballade. This arrangement was intended to provide an 'ungrammatical' version of the first section of this musical piece.The new 'ungrammatical' order was obtained aleatorizing the order of the different units, but avoiding the location of the studied units at the beginning and at the end of the sample, in order to control for primacy and recency effects.

The musical piece was not familiar to the participants of this study.

V.1.2.3 Apparatus

An expressive commercial recording of the Chopin's Ballade No.1, in G Minor, op. 23 was selected [Jonathan Shin'ar, piano (recorded in London, April 1995) 1996, Carlton Classics 30367 00992]. The first section of the Ballade (mm 1 to 36) was extracted and edited using Soundforge 6.1 software. The musical samples were CD recorded, keeping the expressive parameters of the original version.

V.1.2.4 Procedure

An on-line parsing paradigm was used in the experiment to test participants' sensitivity to different degrees of closure. In each trial, subjects listened to a sequence of musical units and were required to assess the degree of closure of each of them as long as they were presented during the silent lapse that mediated between two successive units (2.5 sec in

concatenated conditions 1 and 3, and 10 sec in the *isolated* condition 2). In order to register their responses, participants were provided with a sheet of paper with fourteen horizontal lines that corresponded to the sequence of fourteen units they listened to. While they listened to each unit they assessed the degree of relative closure writing a vertical mark at a certain point of the horizontal line that represented a continuous scale that ranged from *Does not close* on the left to *Does Close* on the right.

The test began with a warm-up session in which subjects were familiarized with the experimental procedure.

Subjects were tested in groups, and sat in front of a stereo CD player that reproduced the musical sequence. The test was administered in a relaxed session that lasted a maximum of 15 minutes (warm-up session included).

V.1.2.5 Design

A factorial design was run, involving one within-subjects factor: *musical unit type* (embedded [5], literally repeated [2] and sequenced [2] musical units, respectively) and two between-subjects factors: *musical experience* (groups 1 and 2) and *condition* (conditions 1, 2 and 3). In order to supply the study with ecological validity, the selection of the grammatical type and the treatment of the musical samples was realized taking into account the order in which the musical units appear in the Ballade; that is, embedded units (see mm. 9.4 to 21.3 in Figure V.1), literally repeated units (see mm. 21.4 to 25.3 in Figure V.1) and sequenced units (see mm. 27 to 30 in Figure V.1).

As each grammatical type of units' combination implies different rationales, three different studies of the perceived closure (one for each grammatical type) were designed: (A) same

melodic material in an *embedded hierarchical context*, (B) same melodic material in a *linear repeated context*, and (C) sequenced melodic material in a *linear sequenced context.*

V.1.3 Results

Several predictions were made:

It was hypothesized that sensitivity to closure would depend on the level of embedding of the musical unit in the hierarchical context (Lerdahl & Jackendoff 1983; Todd 1985; Temperley 2001); so, the more embedded the unit within the hierarchy, the more different the listener's sensitivity to closure of that unit according to whether it is within context (concatenated) or out of context (isolated).

It was assumed that, in linear contexts (as defined above) the locus of the unit contributes to the perceived relative closure; thus, both repeated and sequenced units should be differentially perceived according to their temporal locus in the successive order (Ockelford, 1991).

It was hypothesized that cohesion of context provides the necessary background for assigning the unit coherent meaning in terms of its relative closure. So, closure will be assessed differently according to whether the unit is concatenated (concatenated condition) or isolated (isolated condition).

It was also assumed that grammaticality of the ordered relationships in the constituent organization of the piece influences the perceived closure of the units. So, sensitivity to closure in a given unit should be different when the unit is within a 'grammatical' context than within an 'ungrammatical' context.

As musical tuition provides an increasing awareness of grammatical musical organization, it was expected that the degree of musical experience would have an effect on sensitivity to musical closure. Thus, differential sensitivity to closure should be expected in students with different degrees of musical training.

Each written response marked on the sheet was measured in millimetres measuring the segment formed between the left extreme of the horizontal line and the mark produced by the participant. The scale used ranged between 0 (left extreme) and 100 (right extreme).

Results are presented according to the three different studies of the perceived closure that were performed.

(A) Same melodic material in an *embedded hierarchical context*

A repeated measures ANOVA with *Units* (5) as within-subjects factor and *Condition* (3) and *Musical Experience* (2) as between-subjects factors was run. Results can be observed in Figures V.2 a and V.2 b.

Factor *Musical Units* showed a significant main effect: $F_{[4\text{-}728]} = 154.273$; $p < .000$. Both *Condition* and *Musical Experience* were also significant: $F_{[2\text{-}182]} = 7.869$; $p = .001$ and $F_{[1\text{-}182]} = 24.788$; $p < .000$, respectively. All interactions were also significant ($p < .000$).

Results in Condition 1 for initial students fit to prediction 1: unit a.3 (the most embedded one) was considered as the most "closed".

Similarly, units a.1 and a.2 reflected both the expected results in terms of embedding. Although unit a.5 is more embedded than unit a.4, the lower scores may be due to an effect of the surface harmony: the tonic chord as a secondary dominant. Differences between condition 1 and 2 seem not to be more pronounced.
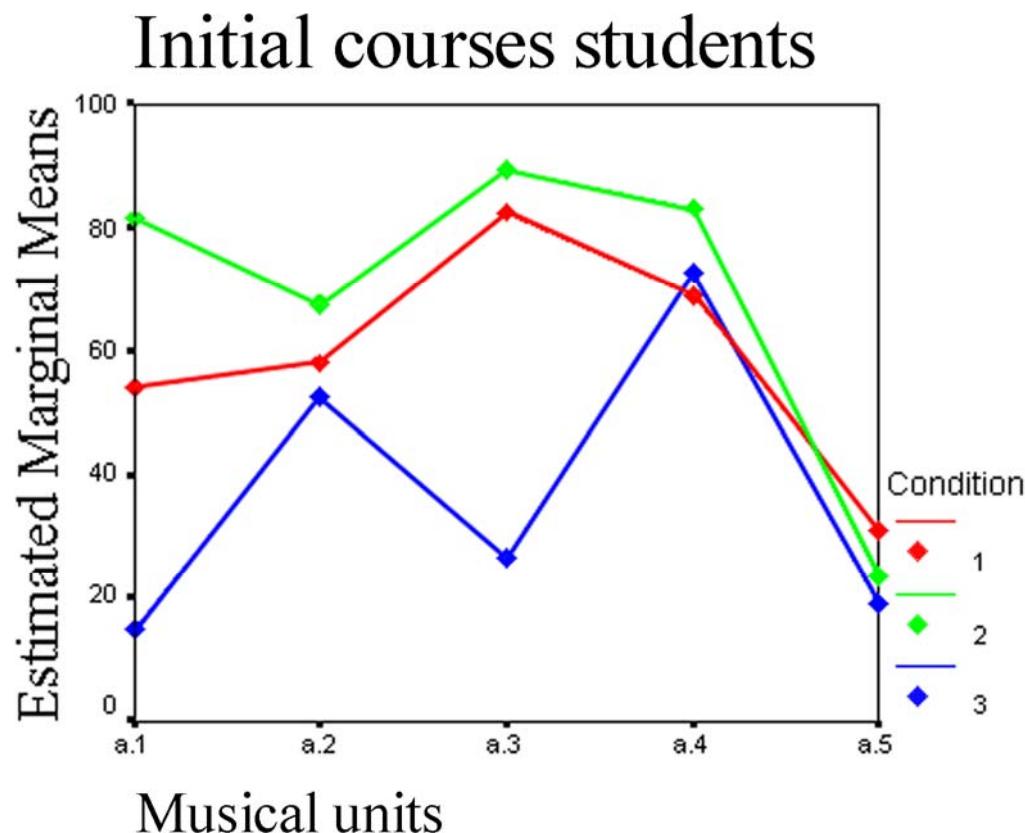
Figure V.2 a Results corresponding to study A (same melodic material in an embedded hierarchical context) for the group of initial courses students

Therefore, it could be said that the span between units in condition 2 (*isolated*) is not sufficient to break the coherence of the grammatical context that the grammatical order provides. On the contrary, results in condition 3 (*'ungrammatical' concatenation*) showed that temporal order is extremely important to build the hierarchy and to perform the attribution of closure to each unit.

Note that more advanced students tend to reduce differences between conditions and units. It seems as if musical training reduced the effect of the musical context.
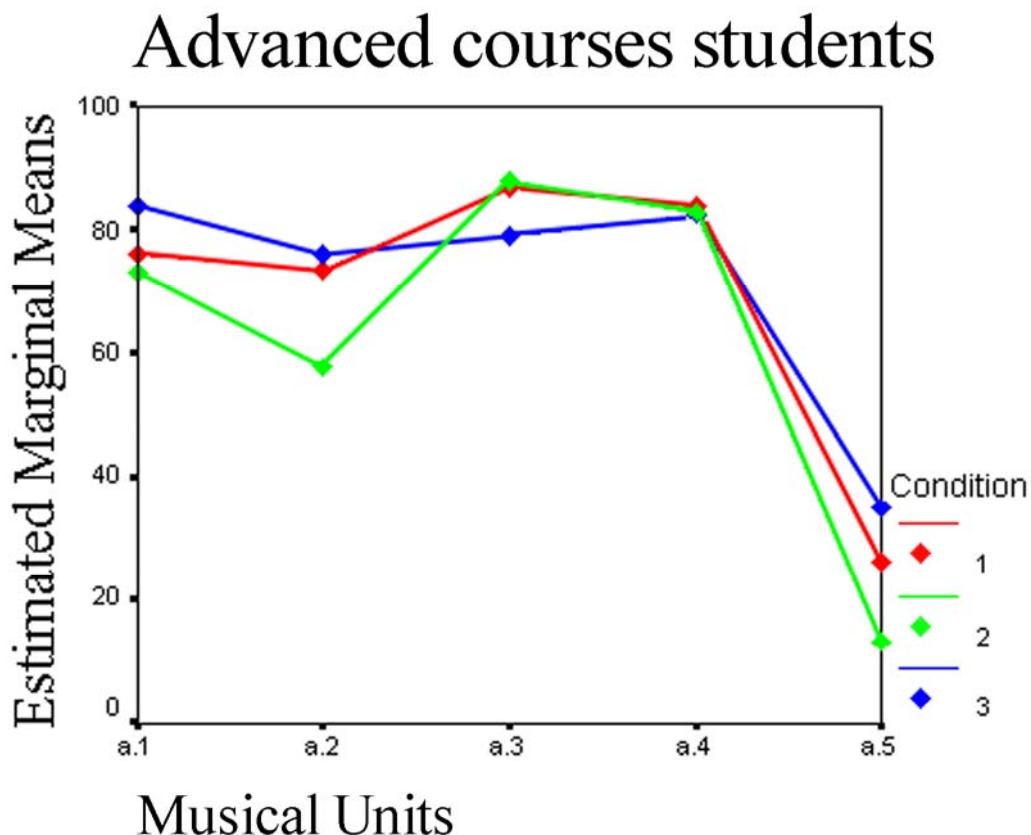
Figure V.2 b Results corresponding to study A (same melodic material in an embedded hierarchical context) for the group of advanced course students

(B) Same melodic material in a *linear repeated context*

A repeated measures ANOVA with *Units* (2) as within-subjects factor and *Condition* (3) and *Musical Experience* (2) as between-subjects factors was run. Results are showed in Figure V.3.

Factor *Musical Units* showed a significant main effect: $F_{[1-182]} = 14.700$; $p = .000$. Participants in both groups considered the second unit (b.2) as less closed than the first unit (b.1).
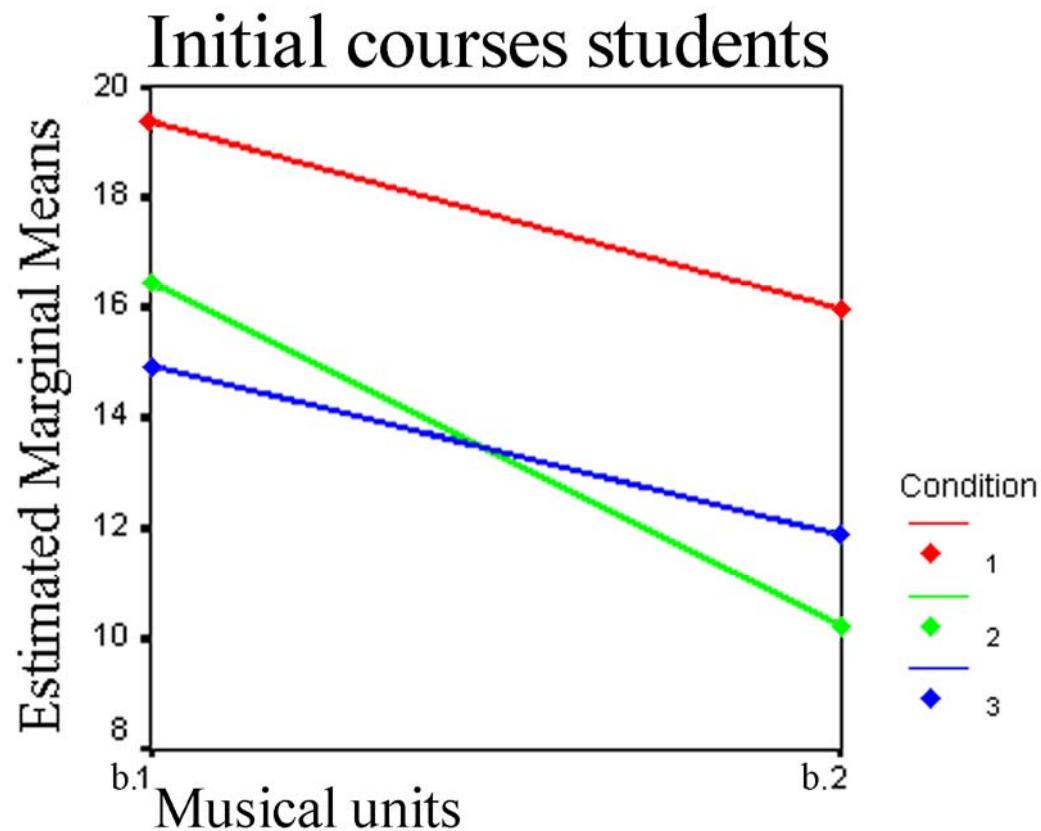
Figure V.3 a Results corresponding to study B (same melodic material in a linear repeateed context) for the group of initial courses students

Results showed differences related to factors *Condition* and *Musical Experience* that, nevertheless, did not reach statistical significance. Predictions about order and/or grammaticality of the context are not clearly fulfilled: in conditions 1 and 2 (concatenated and isolated linear repetition) there is a tendency to react to the locus of the unit in a similar way.

In condition 3, on the contrary, it would be expected that the 'ungrammatical' order generated by the context had an effect in the perceived closure. However, the tendency that arose from the outcomes resembles what was observed in conditions 1 and 2.
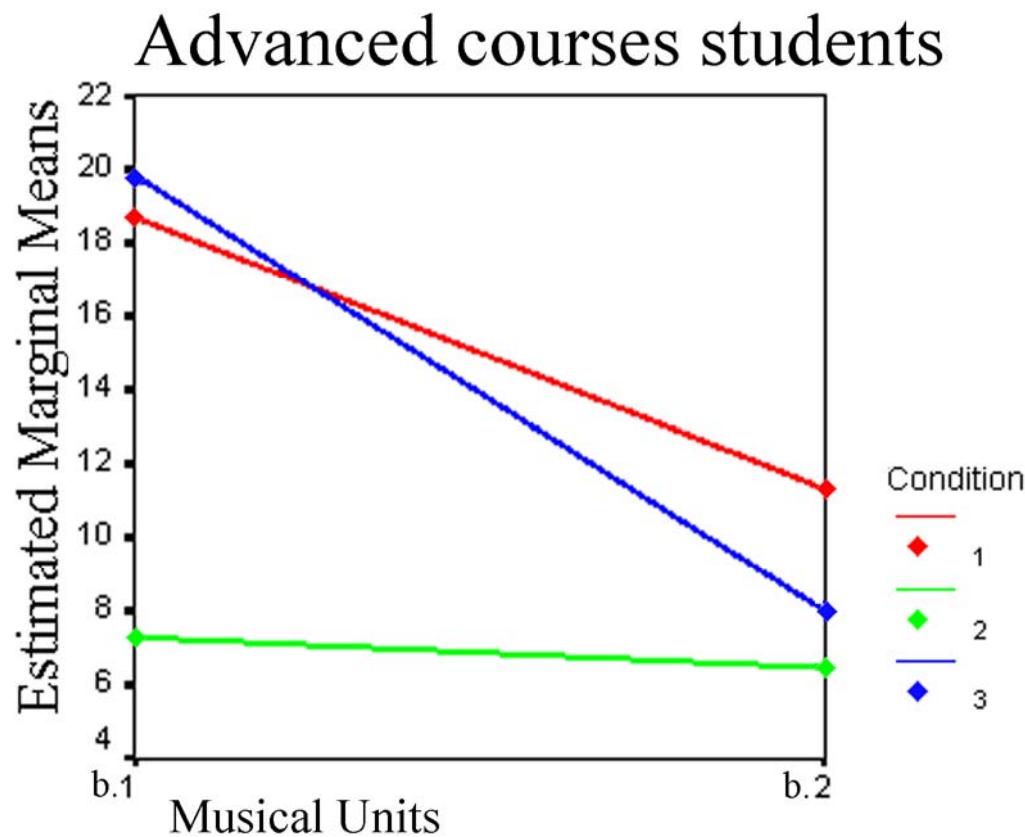
Figure V.3 b Results corresponding to study B (same melodic material in a linear repeateed
context) for the group of advanced courses students

Nevertheless, it is not possible to conclude that 'ungrammatical' arrangements do not

affect sensitivity to closure in linear contexts given that, in the present study, a single

musical piece was used in order to account for ecological conditions to test participants'

music perception. The need to create a natural environment, as close as possible to a real

time listening situation, constrained the number of opportunities available to experience

perceptual closure to only one presentation of each grammatical context. It might be the

case, for example, that the peculiar 'ungrammatical' arrangement created as a result of the

aleatorization of the fourteen units - in which units b.1 and b.2 were included - , could have

prompted the particular outcomes found here. Note that, although in condition 3 these units

170

were randomly ordered, subjects in that condition listened to *only one* 'ungrammatical' order. Besides, it might be possible that, in spite of repetition, attention to textural surface details in unit b. 2 (increment of melodic activity in the bass line) had an effect in the outcomes of this study. This idiosyncratic attentional mode was already observed in the outcomes of study A with respect to voice-leading components of repeated units in an embedded context (see above). Therefore, the methodological constraints of this musical study did not allow knowing more precisely which were the effects that the locus of a repeated unit, or the surface quality of a unit in a repeated context had in perceptual closure; it would be necessary to perform further studies in order to derive more definite conclusions concerning this type of grammatical context.

(C) Sequenced melodic material in a *linear context*

A repeated measures ANOVA with *Units* (2) as within-subjects factor and *Condition* (3) and *Musical Experience* (2) as between-subjects factors was run. Results can be observed in Figure V.4.

Factor *Musical Units* was not significant. *Condition* showed a significant main effect: $F_{[2-180]} = 10.346$; $p < .000$. But *Musical Experience* was not significant. Interactions between *Condition* and *Musical Experience* and *Units* and *Musical Experience* showed marginal effects (F = 4.248; $p < .016$ and F = 5.075; $p < .025$ respectively). Interaction among the three factors also presented a marginal effect (F = 3.871; $p < .023$). This means that although students tended to assign different levels of closure, their attribution was strongly dependent on the condition.
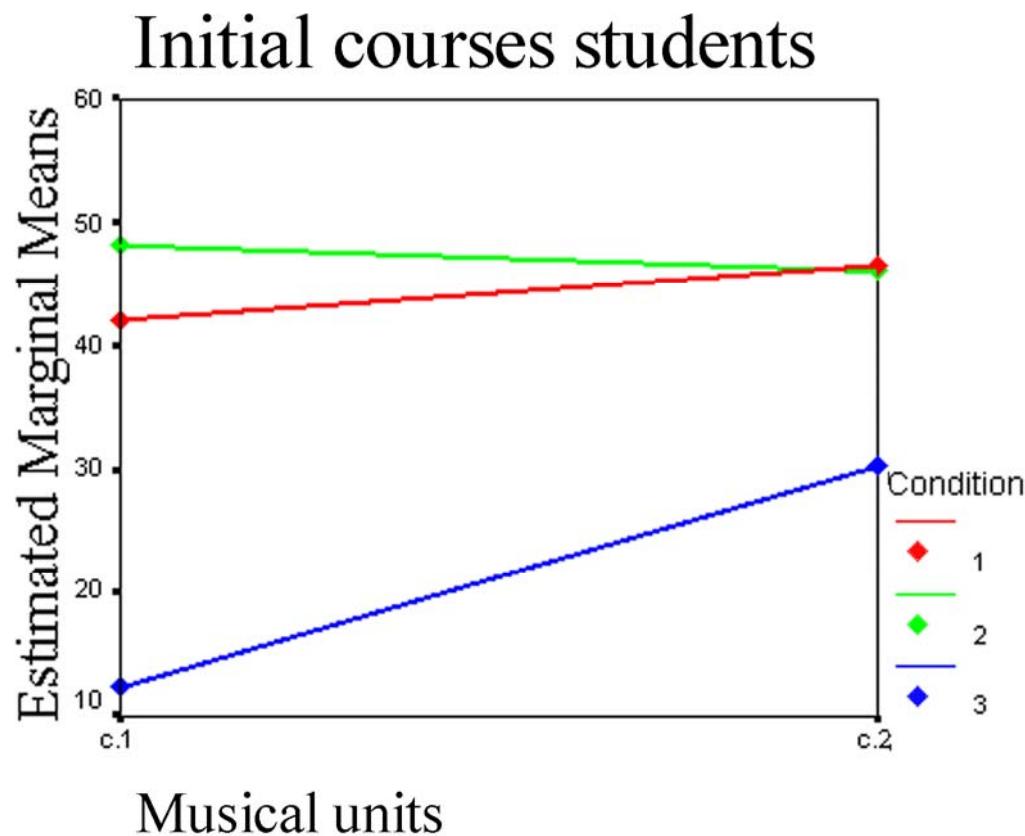
Figure V.4 a Results corresponding to study C (same melodic material in a linear sequenced context) for the group of initial courses students

In condition 1 (concatenated units in sequenced continuation), it could be observed in both groups a tendency to perceive sequenced continuation in such a way as to consider the second unit as more closed than the first one, which, at a local level, is coherent with the characteristic musical construction of the passage. In condition 2 (isolated units in a sequenced continuation), on the contrary, it seems as if the temporal span that separates the units prompted in both groups a tendency to highlight different surface details in music perception, so as to consider the second unit as less closed than the first one.
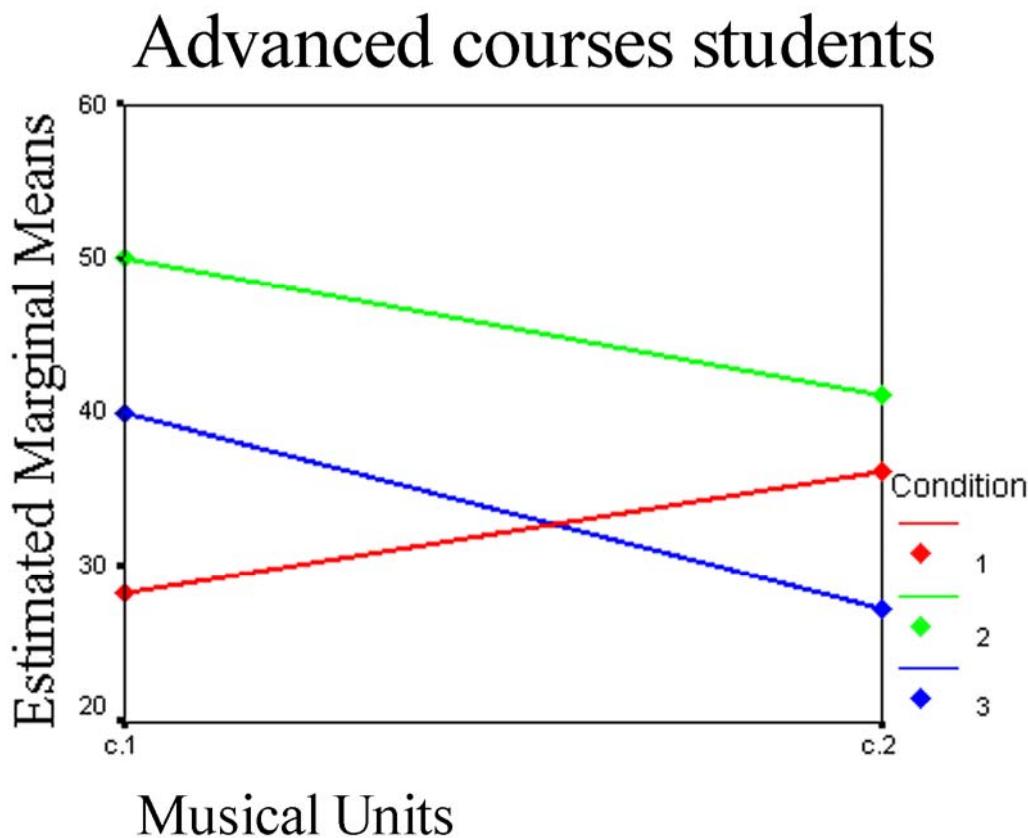
Figure V.4 b Results corresponding to study C (same melodic material in a linear sequenced context) for the group of initial courses students

While in condition 3 (ungrammatical concatenation) the context seems to produce one effect for the initials and the opposite the advanced music students. In other words, both continuation and order of the sequenced linear context seem to be important for the perceived closure,although they would be influencing in a different way. As was noted before concerning results of study B, it would be interesting to test a variety of instances of this particular musical context in order to study further the effect of the above-mentioned musical factors in the perceived closure.

V.1.4 Discussion

The aim of experiment V.1 was to test the listeners' sensitivity to relative degrees of musical closure in constituent organizations. Different constituent contexts in which relativity of closure could possibly operate were investigated. Concurrently, several predictions about (i) the relationships between type of context (hierarchically embedded, linearly sequenced, and linearly repeated), (ii) type of listening condition (concatenated units or isolated units), (iii) grammaticality vs. non-grammaticality of the constituent concatenation, and (iv) effect of musical training in the perception of musical closure were formulated.

In study A, it was found that the degree of embedding in the hierarchy seems to be crucial to estimate the relative closure of meaningful musical fragments. In hierarchically embedded musical contexts, cohesion and grammaticality of the constituent function seem to be related as much to the concatenation of the units within the ordered syntactic sequence as to the musical discourse as a whole: similar results found in conditions 1 (concatenated grammatically arranged) and 2 (isolated grammatically arranged) reveal that subjects with an initial musical experience were able to identify the function of the grammatical units in the embedded structure even when the units were separated from each other by a long temporal lapse, whenever they were listened to it in the "correct" grammatical order.

In study A, condition 2, it is possible to explain the higher estimation of the perceived closure in unit a.1 in terms of the underlying voice-leading: unit a.1 presents a complete linear progression (D-C-B flat-A-G) while unit a.2 appears as a prolongation of the above mentioned G (double neighbour F#-A-G). This analytical feature provides a plausible explanation of the reason why unit a.1 is understood as more closed than unit a.2 when

they are presented in an isolated fashion. In connection with this explanation - and given that for subjects with initial musical training it appears that the context has a more pronounced effect in the experience of closure - in condition 3, the peculiar configuration of results for initial students might be related to the particular *ungrammatical order* of the units' presentation to which they were exposed. Differences between groups in the perceived closure in 'ungrammatically' ordered contexts are equally noticeable in the scores evidenced in condition 3 of study C. These results that emerge from the non-grammatical arrangement highlight the need for further investigation, testing the influence of new un-grammatical contexts on the perception of relative closure; however, accomplishment of this goal is beyond the objective of the present thesis.

The performance of more advanced students on this test, in particular in all what concerns to perceptual closure in embedded contexts, suggests that the systematic musical training favours the use of a more local listening strategy, in which conditions of closure are determined according to the operation of local factors (such as the surface harmony, melodic design, textural details, etc.) rather than by contextual factors (related to the grammatical function of the meaningful units). This idiosyncratic mode of using music cognition might be the result of changes in the students' use of learning mental strategies derived from the curricular focus on topics such as the identification of *atomic* attributes of musical structure (i.e. chords, intervals, rhythmic cells and note spelling, among others) in courses of aural training and music theory. This interesting conclusion might inform the field of music education, highlighting the need of keeping and enhancing the myriad of generic processes that constitute cognitive reference points (Rosch, 1978) in the development of musical ability.

But much more important to the purposes of the present thesis is the fact that current results help to clarify previous discussion about the cognitive status of the experience of closure in tonal music. They show how closure is inherently related to the syntactic contexts in which it is experienced. In embedded contexts, the notion of closure appears clearly as highlighting the relational component of musical grammar. While in linear concatenations it is not as clear the way in which musical structure influences perceptual closure. It seems as if each unit is perceived more as a self-contained entity.

It is also evidenced in this study that some differences in the perceived degree of closure are the result of the interaction between surface and deeper features of the contrapunctal-harmonic organization of the piece (see above in this section).

Furthermore, results confirm previous assumptions according to which embedded repetition is understood differently to linear, coordinated repetition (Horton, 2003).

Finally, results provide behavioural evidence of the cognitive status of *constituent boundaries* in different types of grammatical contexts. Closure, while showing a capacity to account for differences in the experience of the internal organization of constituent structures, validates the cognitive status of boundary locations as focal points in music attending. Thus, it opens the door to investigate the experience of other structural features of tonal music: if constituent boundaries have cognitive reality in standard constituent structures, then it is possible to ask about the constituent status of prolongational structures, testing its cognitive experience at boundary locations. This is the main purpose of the experimental work that follows.

V.2 Experiment V.2: Music attending to close prolongations using a click-detection task

V.2.1. Introduction

The aim of the current study is to seek behavioural evidence of the experience of prolongational structure during on-line attention to music. In the present investigation, the prolongational status of pitch events in a tonal composition is hypothesized as a constituent organization, and its experience as a percept is investigated. Based on the results of experiment V.1 it is possible to say that constituent organizations *have* psychological reality and that constituent boundaries are focal points that are experienced in terms of different degrees of closure. In order to arrive at the formulation of the experimental hypothesis of experiment V.2, the main ideas developed in Chapter IV will be summarized in what follows with the aim of presenting the rationale of this experiment.

V.2.1.1 The perceptual status of the prolongational structure

Some of the main musicological theories contain statements that convey the idea of the underlying musical structure as a percept. For example, there are statements that refer to cognitive assumptions such as "the ear interprets" (Salzer & Schachter, 1969; p.119); "the musical ear perceives" (Salzer & Schachter, 1969; p.123); "the ear connects" (Salzer & Schachter, 1969; p.135), etc.

Salzer & Schachter (1969) describe the theoretical concept of the underlying musical structure based on a perceptual-cognitive assumption:

> "The musical listener relates the immediate event - the passing, neighboring or embellishing
>
> chord - to the governing sonority; in so doing he hears, as it were, two events at the same time

but on different levels of perception. The underlying chord, though not always acoustically

present, functions as a background against which other tones and sonorities unfold" (p. 147).

The experience of the underlying structure is closely related to the idea of mental retention, understood in Schenker's writings as a conceptual "mental" state, by means of which a structural tone is kept active (mentally retained) in mind, while other tones unfold a motion at the surface level, until the next tone of the structural upper voice becomes active (see in chapter I a more comprehensive explanation of the idea of underlying structure and the concept of mental retention).

As underlined before in this thesis, several psychological principles may account for the plausibility of the underlying structure as a percept in music cognition (see a detailed description in Chapter II). However, they have been tested mainly at the note-to-note level of the musical surface.

Some behavioural and developmental evidence on the experience of the underlying musical processing has already been collected (see, for example, Dibben, 1994 and Bigand, 1990 for experimental testing on related issues and Serafine, 1988; Serafine et al., 1989; Martínez & Shifres 1999 a, b, c; 2000; Shifres & Martínez 1999, 2000a; Martínez & Shifres 2000 a, b and Martínez, 2000 a, b; 2001 for experiments closely related to the problem tackled in this thesis). Nevertheless, the difficulty of setting experimental conditions that will incontestably account for the cognitive processing of this construct make it necessary to explore further its psychological status. This is the reason why sensitivity to the prolongational structure needs to be tested in terms of a variety of cognitive processes that the listener is assumed to deploy when s/he experiences it. As suggested in Chapter II, the psychological plausibility of this construct could be increased if data derived from different sources of evidence are collected.

As noted above, it is a cognitive assumption relative to the experience of prolongation that the listener relates immediately concatenated events - passing notes, neighbour notes, or embellishment chords - to a non-immediate sonority that "dominates" them at a higher level of the underlying hierarchy. According to this belief, listening to a given musical passage in prolongational terms will elicit a cognitive process that splits the sound sequence into two events that occur simultaneously, but at different perceptual levels. Due to contextual factors that regulate relationships of relative stability between pitch events, if a tone is contextually stable it will play a more structural role, extending its effect throughout a musical passage until a new structural tone is activated. From a psychological point of view, then, the cognitive status of a prolongational extension could be hypothesized as the result of a process in which the structural tone is phenomenologically extended throughout the prolongational unit, whether it is physically present or not, performing the function of a background over which the other notes are developed. In order to produce this result, temporal and non-temporal processes combine in such a way to provide the necessary framework to interpret incoming information in terms of the unfolding voice-leading. It is, therefore, an assumption regarding the experience of the underlying musical structure that the listener develops sensitivity to those elaborative processes that unfold it in time, because the unfolding of the sequence of events somehow orients the way in which attention to music takes place (see Jones, 1992; Jones & Boltz, 1989).

V.2.1.2 Music attending and the constituent status of the prolongational structure

It was proposed in Chapter IV that some of the constructs employed by syntactic theory to formalize relationships of dominance-subordination between events in hierarchical organizations may be useful in studying the underlying structure in tonal music. In

particular, we referred to constituency and dependency as two structural features that may account for the way the components of a musical piece are organized in hierarchical terms.

According to the theory of prolongation, the constraints that govern an underlying voice-leading arrangement in a musical passage regulate the relationships between pitch events in terms of their relative structural priority. Consequently, different hierarchical statuses can be assigned to the pitches of a prolongational unit.

If the temporal unfolding of the structural make-up of a musical composition "orients" the process of music attention, then information processing will be a function of the surface structure of the composition, and attention will vary according to the structural importance of the musical event on which the listener is focusing. Consequently, decisions about what is being heard will be related to the internal organization of the prolongational units into, at least, two hierarchical orders of events: the constituent head and the successive dependent pitch events.

To attribute prolongation the capacity to organize attending to music is to assign it a function that is similar to the functions constituent units are assigned in linguistic theory. If prolongational structure is assigned a constituent function, by means of dependency an important structural note, the head of the constituent, will be prolonged through other embellishment notes that, as far as they depend on it, will elaborate the first sonority, until another structural note is reached (see a detailed description of this organization in Chapter IV).

If prolongation, understood as a one-to-many relationship between pitch events, is assigned a constituent function, the following assumption can be derived: within a musical passage, the listener will experience certain events of the musical surface, as being related to, and

dependent on, some other non-immediate events. Consequently, it is possible to expect that such a one-to-many relationship in some way models the listener's attending process. If the listener somehow shows sensitivity to this type of one-to-many relationship, s/he would be involved in a contextual situation where two events are experienced simultaneously but at different hierarchical levels. Consequently, the head of the constituent would tend to be experienced as a dominant event, and the following pitch events would be experienced, on the one hand, as attached to the former head and, on the other hand, as eliciting expectations of continuation toward the following constituent head.

V.2.1.3 The assumptions

The following general research question is derived from above:

Does prolongational structure organize the way we attend to tonal music?

Based on evidence that (i) structural components of the input influence attention processing (Jones and Boltz, 1989) and (ii) attention does not parallel the temporal linearity of the input (Fodor and Bever, 1965; Holmes and Forster, 1970; Sloboda and Gregory, 1979; Stoffer, 1985; Kaminska & Mayer, 1993) that is to say, information about what is being heard is not continuously available, but is periodically accessible, enabling the perceptual organization of a stream into coherent fragments, the following assumptions can be formulated: i) Information processing is a function of the surface structure of the musical piece. That is to say, the underlying structure orients the way the listener attends to the incoming information. Therefore, attention will vary according to the structural importance of the musical event the listener is focusing on, and decisions about what is being heard will be related to the constituent structure of the prolongational units.

ii) Given that results of experiment V.1 provided evidence of the cognitive status of constituent boundaries as focal points in constituent organizations, then it is possible to investigate prolongational boundaries in order to look for evidence of their cognitive status during music attending.

iii) The amount of processed information is an informative measure of the prolongational status of pitch events in the constituent unit. Information processing will be maximum within the constituent unit and minimum at constituent boundaries.

iv) Information processing can be measured according to the Subject's Reaction Time (SRT) to the required task. SRT is considered a measure of the amount of information that the subject processes in real time (see chapter II for detailed information about SRT).

V.2.1.4 Hypothesis

In the present experiment a click-detection paradigm is used with the aim of testing the following hypotheses:

1) Click-detection time, and hence SRT, will vary according to the prolongational level of the focal point where the click is located.

2) SRT will be faster for clicks located at prolongational boundaries, and slower for clicks located before such boundaries, within the constituent unit.

V.2.2 Method

V.2.2.1 Subjects

Fifty-eight subjects volunteered to participate in the experiment. Thirty-one were professional musicians and twenty-seven were non-musicians. The musicians had an average age of 29 years (range 20-44 years) and an average musical experience of 15

years. The non-musicians had an average age 24 years, and were university students recruited from social sciences courses. All of them had normal hearing. Musicians were defined as graduate professionals with a degree of music at the university or at the conservatoire, and with at least 10 years of instruction on a musical instrument and a mean number of 6 years of instruction in music theory. Non-musicians had not taken musical lessons, had no experience of playing musical instruments or in music reading.

V.2.2.2 Stimuli

11 fragments of musical pieces belonging to the repertoire of tonal Western art music were used (see Figures V.6-V.10 at the end of this section). All the fragments are musical phrases containing close foreground prolongations, that is to say, particular types of linear elaborations of a single structural note, generally belonging to a triad, that end on the same note before progressing to a subsequent structural note. The reductional analyses of their prolongational structures were taken from the literature on music analysis (Forte & Gilbert 1982; Salzer [1962]-1982; Salzer & Schachter 1969).

The length of each prolongational unit was measured in seconds. The mean prolongational unit length was 4.028 seconds with a standard deviation of 1.310. Thus, the prolongational units used in this experiment fall within the frame of the psychological present (Gabrielsson, 1993).

The last note of each of the prolongational units was established, and the prolongation boundary was defined as the focal point determined by the inter-onset interval between the end of the last note of the prolongational unit and the beginning of the following note, that is, the first note of the next constituent unit. The mean inter-onset interval of the 11 fragments at the boundary area was 365ms.

The experimental treatment of the stimulus was performed as follows:

1) In each musical fragment two click-positions were established.

a) Click-position at the prolongation boundary: located in the middle of the inter-onset interval between the last note of the prolongation and the following note (see Figure V.5. a).

b) Click-position before the prolongation boundary: located 1 second before the last note of the prolongational unit (see Figure V.5.b).

Figure V.5.a Mozart, Rondó K 494. Musical fragment with click position at the prolongational boundary and last note of the prolongational unit at weak metrical position

Figure V.5.b Mozart, Rondó K 494. Musical fragment with click position before the prolongational boundary and last note of the prolongational unit at weak metrical position

2) On the basis of previous studies of musical structure, in which an effect of the metrical salience of the structural event was found (see description of the Serafine et al. experiment in Chapter III), it was hypothesized that other factors, such as the relative beats' salience in the metrical structure, might affect SRTs. For example, if more salient beats were very close to the clicks' locations, they might elicit faster RTs. Therefore, the experiment included a condition in which the metrical positions of the notes at the prolongational boundary were changed in order to monitor the influence of the metrical component.

Alteration of the musical passages so as manipulate the relationships between the metrical structure and the prolongational structure inevitably introduced certain changes into the music. These changes were confined to the level of the musical surface, and care was taken to ensure that they would not be likely to manifest themselves at middleground or background levels of structure in any conceivable Schenkerian reduction of the passages used. It can reasonably be claimed that, in reference to Schenkerian analytic conceptions of the passages, prolongational structure was not altered by the changes that were introduced. While it is possible - indeed, likely - that the passages would be perceived as different by virtue of the alterations, it is contended that these perceived alterations arose by virtue of alterations to features that are likely to have been neutral in respect of the prolongational structures of the passages.

Two different metrical positions were used at the prolongation boundary:

a) Metrical position 1: last note of the prolongational unit at a weak metrical position (see Figure V.5.c.1).

Figure V.5.c.1 Mozart, Rondó K 494. Musical fragment with click position at the prolongational boundary and last note of the prolongational unit at weak metrical position

b) Metrical position 2: last note of prolongational unit at a strong metrical unit (see Figure V.5.c.2).

Figure V.5.c.2 Mozart, Rondó K 494. Musical fragment with click position at the prolongational boundary and last note of the prolongational unit at strong metrical position

3) Finally, melodies without clicks were included, in order to detect false alarms (De Witt and Samuel, 1990; Stoffer, 1985) and to avoid serial position effects (Holmes and Forster, 1970), due to anticipations of the click locations by learning during the test.

Figures V.6 - V.11 show the musical samples used in the experiment.

Figure V.6. Musical samples used in the experiment. a: musical fragment with click position at the prolongational boundary and last note at weak metrical position. b: musical fragment with click position before the prolongational boundary and last note at weak metrical position. c: musical fragment with click position at the prolongational boundary and last note at strong metrical position. d: reductional analyses of the musical fragments.(Mozart, Fantasy in d minor KV 397, Adagio, extracted from Salzer, [1962] - 1982, ex. 207; Haendel, Sonata in e m, II, extracted from Forte & Gilbert, 1982, ex.23).

Figure V.7. Musical samples used in the experiment. a: musical fragment with click position at the prolongational boundary and last note at weak metrical position. b: musical fragment with click position before the prolongational boundary and last note at weak metrical position. c: musical fragment with click position at the prolongational boundary and last note at strong metrical position. d: reductional analyses of the musical fragments. (Mozart, Rondó K 494, extracted from Salzer, [1962] - 1982, ex. 332; Bach, Little Prelude in C m., extracted from Salzer, [1962] - 1982, ex. 249).

Figure V.8. Musical samples used in the experiment. a: musical fragment with click position at the prolongational boundary and last note at weak metrical position. b: musical fragment with click position before the prolongational boundary and last note at weak metrical position. c: musical fragment with click position at the prolongational boundary and last note at strong metrical position. d: reductional analyses of the musical fragments. (Chopin, Mazurca op. 41 no. 4; extracted from Salzer, [1962] - 1982, ex. 250; Beethoven, Piano Sonata Op. 27 no. 2; extracted from Salzer, [1962] - 1982, ex. 375).

Figure V.9. Musical samples used in the experiment. a: musical fragment with click position at the prolongational boundary and last note at weak metrical position. b: musical fragment with click position before the prolongational boundary and last note at weak metrical position. c: musical fragment with click position at the prolongational boundary and last note at strong metrical position. d: reductional analyses of the musical fragments. (Beethoven, Piano sonata in C m. Op.13, II, mm 37-40; extracted from Salzer, [1962] - 1982, ex. 383; Schubert, Piano sonata in B flat. D. 960; extracted from Salzer & Schachter [1969]-1989, ex. 7/79).

Figure V.10. Musical samples used in the experiment. a: musical fragment with click position at the prolongational boundary and last note at weak metrical position. b: musical fragment with click position before the prolongational boundary and last note at weak metrical position. c: musical fragment with click position at the prolongational boundary and last note at strong metrical position. d: reductional analyses of the musical fragments. (Mozart, Rondó K 494, extracted from Salzer, [1962] - 1982, ex. 332; Schubert, Sonata en Sib. D. 960. Extracted from Salzer & Schachter [1969]-1989, ex. 7/79).
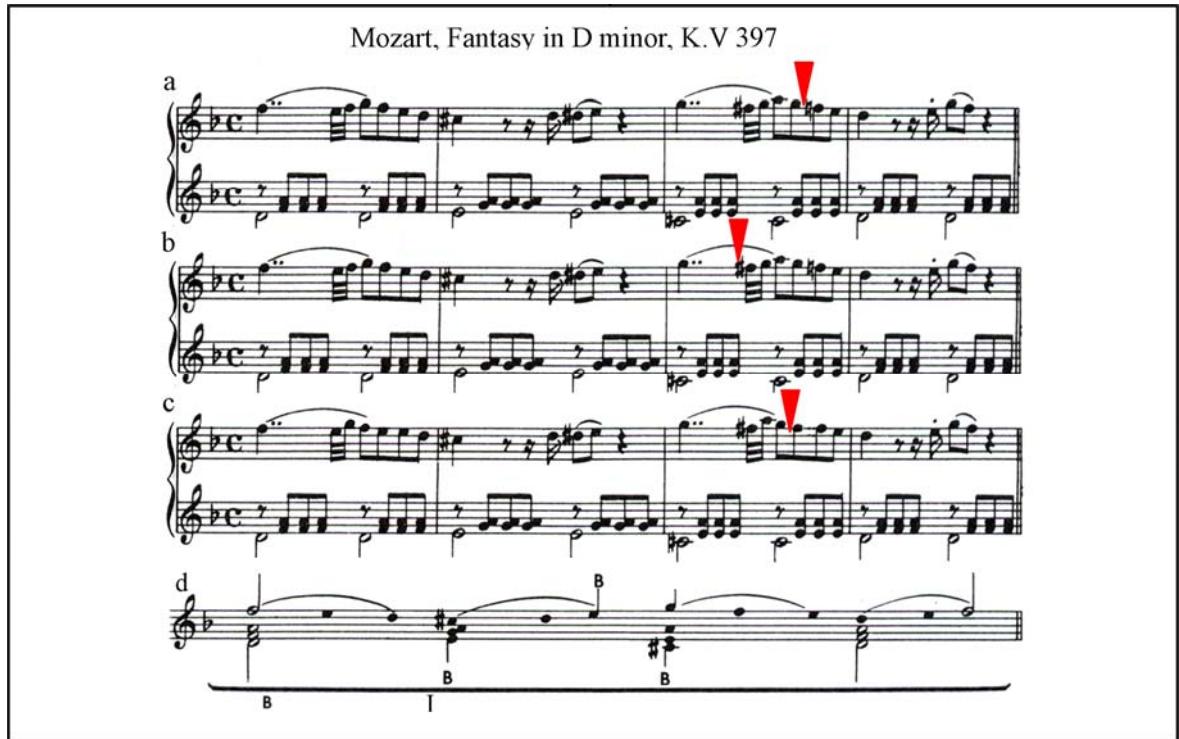
Figure V.11. Musical samples used in the experiment. a: musical fragment with click position at the prolongational boundary and last note at weak metrical position. b: musical fragment with click position before the prolongational boundary and last note at weak metrical position. c: musical fragment with click position at the prolongational boundary and last note at strong metrical position. d: reductional analyses of the musical fragments. ((Mozart, Fantasy in d minor KV 397, Adagio, extracted from Salzer, [1962] - 1982, ex. 207).

V.2.2.3 Apparatus

The melodies were sequenced into a MIDI sequencer (Cakewalk 9.0). They were reproduced using the timbre Piano 1 belonging to the bank 0-Roland GS capital tones from the sound box Roland SC-55 mkII. Parametrization of the musical samples was accomplished in order to keep track of expressive components. Tones were generated by numerical writing of their inter-onset intervals. The velocity of the accompaniment was controlled at a generally lower level sonority than the level of the melody equal in all the samples. Each pitch was quantized and normalized in order to obtain an audio digital track, with a sound level equal to the click, measured in decibels. The click was a digitally

generated sine wave, with a duration of 7ms at a frequency of 1000 Hz. Subjects listened to the sound sequence via headphones (Senheiser model HD 435). The sound sequence was reproduced at a band frequency of 44,100 Hz, 16 bit rate in two channels. The musical fragment was presented in both channels but 4/5 of the signal went through the left channel and 1/5 through the right channel, resulting in the presentation of the left channel 14 dB higher than the right channel. The click signal was presented through the right channel. The inclusion of left channel signal in the right channel had the aim of avoiding aural fatigue. Data from the subject's answers were recorded using Soundforge 4.0 software.

V.2.2.4 Procedure

The experiment began with a familiarization session in which subjects listened to the eleven musical fragments that formed the original samples of the test, repeated three subsequent times each. They were required to listen to the pieces as naturally as possible trying to keep them in mind. The familiarization session was followed by a short warm-up session in which subjects received instructions as to how to proceed with the required task. They rehearsed the click-detection task in the following way: the sound of the click was presented first isolated, and then superimposed to a rehearsal melody. Subjects tried the click-detection procedure three times. They were requested to answer as quickly as possible and precisely, avoiding click-detection errors and key-pressing errors. Once the warm-up session was completed each subject was administered the test in two listening sessions with a ten minute interval of rest between them. In that interval the subject filled in a musical experience questionnaire.

The experiment required a divided attention task with musical listening as the primary task and click-detection as the secondary task. Instructions provided to subjects required them to run two tasks simultaneously:

i) to listen carefully to the fragment and to press a key as soon as they detect the click, and

ii) to recognize if the musical fragment currently heard was the same as any of the fragments heard in the familiarization session.

The recognition task had two aims: i) to focus subjects' attention on the musical structure and ii) to distract subjects to the unique attention to the click occurrence, in order to control for a bias in their responses, reported in previous studies (see Chapter IV).

Subjects were tested individually.

V.2.2.5 Experimental Design

The test employed 66 trials organized in the following way:

-11 trials with click-position and strong metrical position at the prolongation boundary

-11 trials with click-position and weak metrical position at the prolongation boundary

-22 trials with click-position before the prolongation boundary and

-22 trials without click

Trials were randomized, and several orders of presentation were produced for administration to the different subjects. Overall, the test lasted one hour but it was divided into two sessions of 30 minutes each, with 10 minutes of rest between sessions.

V.2.3 Results

Subjects' responses were classified into hits, misses and false alarms. Only 2 false alarms and 5 misses were found. Given their insignificant number they were interpreted as missing values in subsequent analyses.

The RT was calculated in milliseconds as a result of the inter-onset difference between the temporal location of the subject's response and the temporal location of the click. The means of the SRT at both click positions were obtained.

The following analyses of results were performed:

1a. Analysis of results of the click-detection task

The means of the responses at the two click positions were compared. An 11 musical examples x 2 click-positions x 2 musical experience repeated measures ANOVA was run (see Figure V.12). Within-subjects factors were Musical examples and Click position (before and at the boundary); between-subjects factor was Musical experience (musicians and non-musicians).

Within-subjects factor Musical examples was not significant, meaning that there were no differences in the SRT's pattern of responses between the different musical fragments. Therefore, data were collapsed to run the following calculations.

Within-subjects factor Click position was significant (F $_{[1, 56]}$ =30.160; p<.000). SRTs of both musicians and non-musicians were faster in the click position at the boundary and slower in the click position before the boundary.

Within-subjects Factor Click Position: $F_{[1,56]}=30.160$; *p*< .000

Between-subjects Factor Musical Experience: $F_{[1,56]}=6,135$; *p*< .016
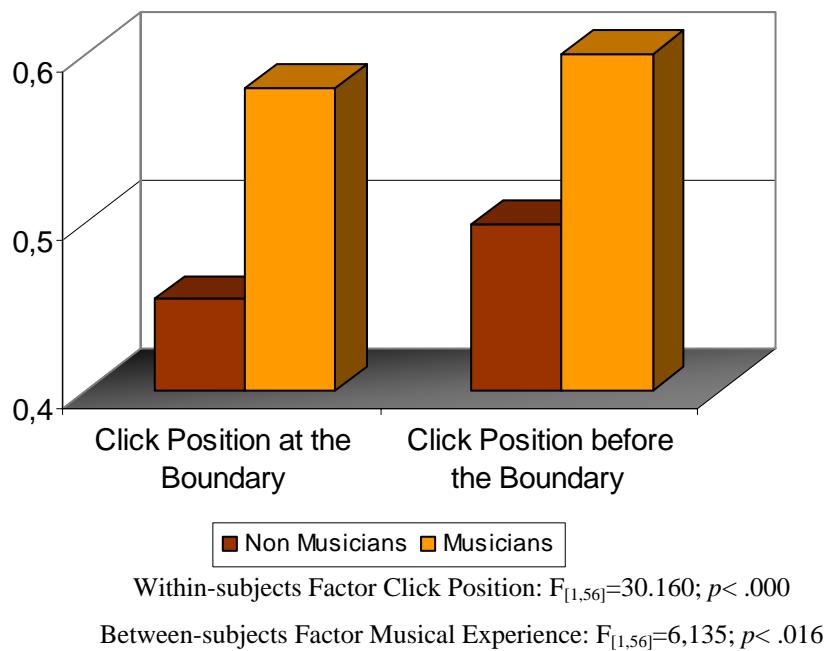
Figure V.12 Comparison between means of SRT for click positions at and before the
prolongational boundary.

Between-subjects factor Musical experience was also significant (F $[1, 56]$ = 6,135; p<

.016). Curiously, non-musicians were faster than musicians. This result will be interpreted

later.

Interaction between Condition and Musical experience was not significant. In spite of the

differences in SRT, both groups showed the same pattern of behaviour at both click

positions. Clicks at boundary locations always elicited a quicker response, independently

of the particular musical example.

1b. Analysis of results of the change of metrical position at the prolongational boundary.

To study the effect of the metrical position at the prolongational boundary, only

participants' responses to musical samples with clicks at boundary locations and changes

in the metrical position were analysed (see description of changes in the metrical position

in section V.2.2.2. above).

The means of the SRTs at metrical positions 1 and 2 were compared. A repeated measures ANOVA did not show significant differences. Therefore, the metrical position at the prolongational boundary had no effect on the subject's sensitivity to clicks located at that focal point.

1c. Analysis of results of the recognition task.

The accuracy of melody recognition between musicians and non-musicians was compared. Correct and incorrect responses were counted. We can see in Table V.1 that, as expected, musicians were more accurate than non musicians (χ^2=296.06; p<.000).

Responses	Musicians	Non-musicians	Total
Correct	2.100	1.287	3.387
Incorrect	628	1.089	1.717
Total	2.728	2.376	5.104

$\chi2$=296.06; *p*< .000

TableV.1 Comparison between musicians and non-musicians in the recognition task.

V.2.4 Conclusions and Discussion

As a result of this experimental work:

i) Behavioural evidence of the listener's sensitivity to prolongation during music attending was found. SRTs were significantly different according to whether clicks were located within the constituent phrases or at their boundaries.

ii) Results provide evidence of the boundary of prolongation as a focal point in music attending, contributing to support its psychological reality.

iii) The way subjects react when they are processing musical information at different focal points of a prolongational unit is in agreement with the idea of *prolongation as a constituent percept*.

iv) SRTs also account for a sense of fulfilment experienced at the focal point of the prolongational boundary, which confirms previous evidence concerning the tendency of structural units to preserve their integrity by resisting interruptions (Fodor & Bever, 1965). It is in this sense that prolongation can be understood as a constituent function of musical structure.

v) The click-detection technique has proved to work well with prolongation, validating the idea of constituency of this construct.

vi) In spite of the similar behaviour demonstrated by musicians and non-musicians in the processing of the two focal points at the constituent units, differences in their patterns of behaviour were found in other respects. Musicians were more accurate and slower than non-musicians. A more precise recognition activity entails paying more attention to the surface details while listening to the unfolding of the piece; this, in turn, makes the processing load heavier. The qualitative difference in the participants' aural analysis might account for the delays found in their RTs. Nevertheless, the resulting association between recognition accuracy and speed of processing deserves further investigation.

vii) Evidence of the incidence of metrical factors in SRT was not found. This result supports theoretical interpretations that consider metrical and linear aspects of musical structure as independent components. The collected evidence seems to indicate that, during music attending to the musical fragments, participants focused more on the continuity of the linear aspects of pitch organization (Schenker, [1935]-1979) than on a parsing

procedure that attended to the time-span organization of rhythmic-metric information (Lerdahl & Jackendoff, 1983).

As was noted above, the musical experience of temporality is in part connected to the idea of filling in temporal intervals (Jones and Boltz, 1989). In tonal music, the unfolding of a sequence of pitch events throghout a given constituent generates expectations of the continuation of events at different focal points of such units, and of a relative sense of fulfilment at its boundary. The click functions as a structural marker: either it conflicts with the ongoing pitch sequence, interrupting it out of the boundary location, or it does not, because a level of fulfilment has been reached at the boundary location. Detection of this signal, measured as the subject's reaction time, is related to particularities of the linear information processing in real time situations. In this experiment, it was expected that participants' differences in their reaction times when clicks were detected at different focal points of the constituent unit could account for differences in the structural statuses of the constituent's events. RTs were in fact different according to whether the signal was located within the prolongational unit or at its boundary. The observed variations in RT can therefore be attributed to the listener's sensitivity to the structural hierarchy of the attribute that is being unconsciously monitored at the click-detection momentum.

The experimental isolation of tonal and metrical factors by means of the manipulation of the note's metrical position at the boundary location did not show statistically significant differences; according to these results, it was concluded that, at least in these musical samples, metrical factors had no effect on SRTs. Although substitutions of the constituent's events might produce alterations in certain aspects of the constituent's intrinsic organization and may elicit, as a consequence, variations in the listener's response, the fact that the change in the metrical position at the boundary location did not

affect RTs supports the assumption that listeners focus their attention more on the tonal pitch characteristics rather than on the metrical components of the constituent organization when they are processing the voice-leading component of musical structure. This apparent independence of pitch and metrical components which occur when prolongational aspects of the unfolding tonal structure are mentally processed support the assumptions of some musicological theories about the primacy of the pitch component in the cognitive representation of underlying levels of tonal structure.

Summarizing, the aim of the present study was to estimate the participants' analytical perception at boundary locations of surface reduction prolongations and to see if it was possible to ascribe the idea of constituency to the composing-out of prolongational structures. The consistency found in participants' responses indicates that the tonal content of a musical composition is not arbitrary, but is at least in part determined by the prolongational functions of its constituent organization, and by the way it is syntactically realized. If this is the case, the final note of a prolongational unit frames the constituent before progressing to the following constituent unit, delimiting the prolongational meaning at a given level of the underlying hierarchy.

V.2.5 Further replications of the click detection experiment

Additional experimental evidence of sensitivity to prolongational structures in a click-detection experimental context has been found elsewhere. An experiment involving children in a click-detection activity brings further evidence of sensitivity to prolongation in music attending in childhood (Shifres & Martínez, 2002). In this study, the assumptions and stimuli were similar to those of study V.2. Thirty children (6-14 years) divided into three age groups took part in the experiment. As in study V.2, results showed faster RTs

for clicks at boundary positions (F [1 - 47] = 5,447; *p* < .025). Younger children seemed to have slower RTs than older children. As in study V.2, metrical positions at prolongational boundaries were also a non-significant factor, although a tendency for the youngest children to be sensitive to it was observed (F = 2.179; *p* < .032). These results provide further ontogenetic evidence of sensitivity to prolongation. The finding is congruent with previous results obtained when a similarity judgment task was employed (Martínez & Shifres, 1999b) adding developmental evidence for the acquisition of structural hearing during childhood (see Chapter III for a detailed account of previous research on the experience of prolongational structure).

V.3 Experiment V.3: Music attending to open prolongations using a click-detection task

A further experimental study was run at the Centre for Music and Science of the Faculty of Music of Cambridge University with musicians and non-musicians as participants. The study aimed at replicating results of study V.2, testing listeners' sensitivity to prolongational structures that represented *open* prolongations during music attending. Seven musical phrases belonging to the repertoire of tonal Western art music were used. They presented a voice-leading unfolding of the notes of the fundamental line using suffix and/or affix open prolongations, a particular type of voice-leading process in which a structural tone of the fundamental line is elaborated through embellishment tones that are placed before and/or after the structural tone (see Forte & Gilbert [1982] – 1992; Cadwallader and Gagne, 1998; Larson, 1997). The experiment was designed to test a constituent function of these open prolongations during an aural perception activity that involved attention to music. If differences in attending according to the structural

component that was being tested occurred, then the reaction time for clicks placed within a boundary between two successive units would be faster than for those placed one second before the boundary of a given unit. The hypothesis that the salience of metrical units would strongly affect the results was nullified by results of experiment V.2 and hence rhythmical significance was ignored as a variable.

V.3.1 Method

A click-detection paradigm was used with the purpose of assessing the listener's attentive response to focal points of open prolongations. A constituent function, analogous to that of syntactic units in speech, was also assigned to these hierarchical units. The experimental procedure replicated that of experiment V.2: clicks were superimposed to musical stimuli and SRT was measured, under the assumption that: i) information processing is a function of the hierarchical structure of the musical piece, ii) information processing is maximum within the constituent unit and minimum at the unit's boundary and iii) information processing may be assessed in terms of SRT, given that attention varies according to the structural importance of the focal point on which the click is placed.

V.3.1.1 Subjects

25 subjects volunteered to participate in the experiment. 6 non-musicians (3 male, 3 female), and 16 musicians (10 male, 6 female) residents at the Cambridge area participated in the experiment.[1] The criteria for the two categories were as follows:

---

[1] We thank Nicky Olle, one of the undergraduate students at the Music Perception and Performance course at the Faculty of Music, Cambridge University, for helping with the recording of the musical samples and the administration of the test.

A Non-musician (NM) was categorised as somebody who had never had individual tuition on any instrument, and who could not read music.

A Musician (M) is a performance-based musician, participating in practical music-making at least once a week, who was or had recently been either a member of Cambridge University Chamber Orchestra, The National Youth Orchestra of Great Britain, held a University Instrumental Award, or attended a London Music Conservatoire.

V.3.1.2 Stimuli

Commercial versions of seven musical fragments of the Western Art Music repertoire were used as musical samples in the experiment. The selection fulfilled the aim of provinding increasingly natural listening conditions of experimental testing. Musical fragments presented voice-leading unfoldings of the notes of the fundamental line using prefix and/or suffix open prolongations (see Figure V.13.a-g).

Two click positions were tested: i) Click-position in the structural boundary: located in the middle of the inter onset interval between the last note of a structural unit and the first note of the following unit and ii) Click-position before the structural boundary: located 1 second before the unit's boundary.

V.3.1.3 Apparatus

The experiment was produced and run in the CMS of the Faculty of Music of Cambridge University using the technological resources available. Emac itunes was used for transferring the CD tracks to audio files. The click was a modified square wave signal generated in Praat, to avoid confusion with any instrument or pitch.

Figure V.13.a-g. Extracts of the musical fragments used in experiment V.3. Prefixes and suffixes are marked in brackets. Clicks located at suffixes and/or prefixes boundaries appear in red.

It was played into the left channel only, with 75% of the total wave in the right channel, and 25% in the left, through a Digidesign Focuswrite Mbox. This avoids audio fatigue affecting the subjects' responses. Responses were obtained using Digital Performer Version 4.1 coming from a Yamaha DX11 keyboard. Subjects used DT 231 Pro Headphones, Beyerdynamic to listen to the musical samples.

V.3.1.4 Procedure

The procedure was almost the same as that followed in experiment V.2. First, each melodic fragment was played twice in the pattern AA, BB, CC, etc. without clicks, to familiarise the subjects with the music. The subjects were instructed to listen carefully to the music in order to internalise it. In this session, there were the seven samples that would be used afterwards in the test, and 5 "dummy" extracts inserted so that even if musicians recognised all the pieces, since they would not hear them all in the latter stage of the experiment, this would show doubt in their mind, preventing complacency and ensuring that their concentration level was high.

Seven of the melodic fragments were then heard twice, in a random order, with a click inserted before and during the salient boundary in each case. During these playings, the subjects were required to perform two tasks: to hear the click, and press a button on a keyboard (marked with a blue sticker) as quickly as possible after hearing it, and to determine whether they heard the extract in the initial familiarization session (there was a red and a green sticker on 2 adjacent keys in a different register of the keyboard to the blue sticker – green=yes, red=no). This choice was made during a gap (approx. 5s) between each playing. The use of a keyboard meant that the set-up could be mirrored for right- and left-handed people. The function of the recognition task ensured that they listened to the

music, and didn't focus exclusively on the click, as this would mean that their reaction time responses would be biased for attending uniquely to the signal.

The experiment was run at the CMS of the Faculty of Music at Cambridge University. Participants were tested individually.

V.3.1.5 Experimental Design

Thus, the test contains 14 trials:

7 of each extract with the click 1s before the boundary.

7 of each extract with the click in the boundary.

V.3.2 Results and Discussion

The means of the SRT at both click positions were compared. A 7 Musical Example factor x 2 Condition factor (click position before and at the boundary) x 2 Musical Experience factor (musicians/non-musicians) repeated measures ANOVA was run (see Figure V.14).

Condition factor was significant (F $_{[1, 20]}$ =11.301; p<.003). SRTs were faster in the click position at the boundary and slower in the click position before the boundary. SRTs are significantly different according to the structural importance of the focal points of the constituent's phrase structures. Musical Example factor was also significant (F $_{[6,15]}$ =9.028; p<.000). This result informs about potential differences in the composing-out of the prolongational units at deeper levels of the underlying hierarchy, which deserves further analysis. Factor Musical Experience was not significant. Both musicians and non-musicians behave in a similar way when they paid attention to the underlying structure of music. This finding is consistent with the findings of a large number of experiments that inform about the absence of significant differences between both populations when

subjects are required to perform tasks that involve this kind of basic musical knowledge (Bigand, 1990; 1994).

Click detection Results
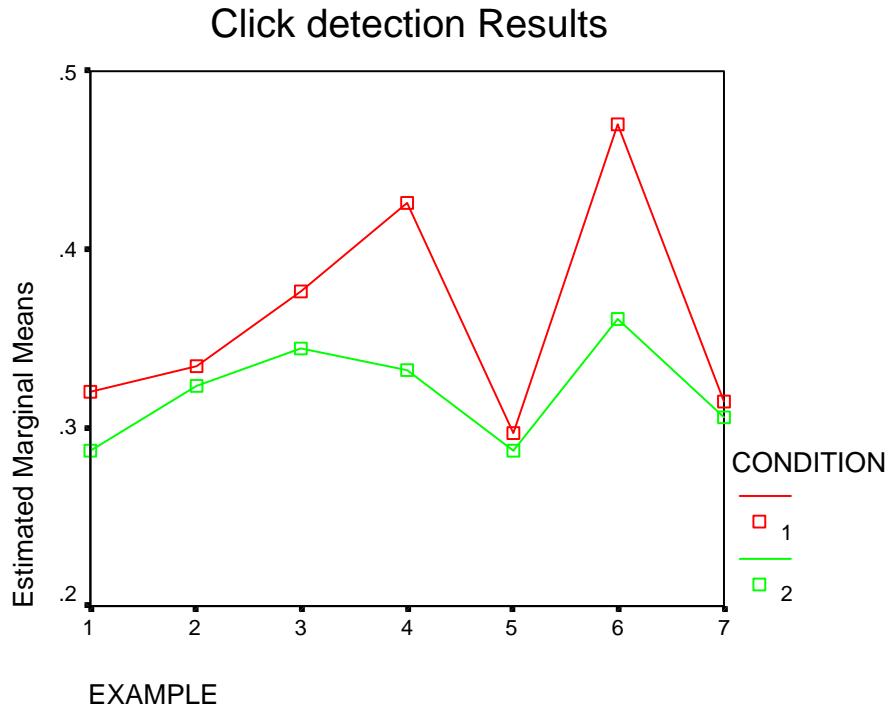
Figure V.14 Click detection results with musicians and non-musicians merged in a single group.

Interaction between factors was not significant either, suggesting that always the boundary prompts a quicker response, independently of the particular musical example.

Therefore, results provide further behavioural evidence of the listener's sensitivity to the underlying structure during music attending. Results also contribute to further support the psychological evidence of the construct boundary of prolongation-progression as a focal point in music attending of the hierarchical music structure. The way in which subjects react when they are processing the information at different focal points of the prolongational unit, is in agreement with the idea of the underlying structure as a percept. SRT, also, accounts for a sense of fulfilment experienced at the focal point of the structural boundary, which is in accordance with previous evidence concerning the tendency of a

structural unit to preserve its integrity by resisting interruptions (Fodor & Bever, 1965). It is in this sense that a prolongational unit could be understood as a constituent function of musical structure. Click detection technique has proved once more to work well in this experimental setting, adding evidence in support of the idea of constituency as a functional feature of the underlying musical structure.

V.4 Experiment V.4: Music attending to close prolongations using a click-location task

V.4.1 The validity of the click-location technique

Experiment V.4 had as its main purpose that of obtaining further evidence of the cognitive status of constituent units, testing the participants' sensitivity to prolongational structures of musical fragments during a click-location task. Previous investigations about speech processing demonstrated a click migration effect, that is to say, a tendency to perceptually migrate the location of clicks toward the boundaries between structural groupings. These studies also suggested that the quantity and direction of the click migration appeared to be a function of the hierarchical order of the phrase boundary where the click was located: the higher the number of superimposed units at a given click position, the higher the listener's tendency to migrate the click to the common boundary of those formal units. Besides, click location accuracy seemed to be higher at constituent boundaries than inside the constituent unit, which may explain variations in the amount of processing load within the constituent unit in a similar way to what was observed in studies that use click-detection tasks.

As reviewed in Chapter IV, click-detection and click-location are two of the main experimental paradigms used to test the cognitive processing of constituent aspects of speech and music. The results of experiments V.2 and V.3, in which the constituent status

of prolongational structures was investigated, showed significant differences in SRTs to structurally different focal points of tonal fragments, using a click-detection technique. Click-detection methodology was successful in providing a reliable measure of the participant's sensitivity to the attribute investigated. If prolongational structure exhibits a constituent conformation when its cognitive status is investigated with one of the paradigms mentioned, it is expected that the alternative methodology of testing will provide further evidence to enhance its psychological status. Thus, experiment V 4 aimed at exploring further sensitivity to prolongation using a click-location task.

However, as noted earlier (see Chapter IV), several shortcomings emerged when the click-location methodology was applied. The results revealed some differences between speech and music processing with respect to the degree of anticipation in the perception of clicks: in speech, clicks tend to be perceived earlier than in music; in music, late responses are much more frequent than early ones. It is not clear whether this difference in the latency of perception is the consequence of differences in the intrinsic configuration of speech and music, or if they are due to particularities of the cognitive processing during the use of this methodology.

In general, music experiments appear to show a click migration effect: results seem to indicate that clicks are located after their real occurrence, except for clicks that immediately follow the phrase boundaries, which are migrated backwards.

But it also emerges from the analysis of those results that click migration is affected differently according to the structural phrase marker that is being manipulated: for example, harmonic markers seem to produce a migration effect when clicks occur after phrase boundaries, indicating that perhaps participants are unable to anticipate those boundaries on the basis of harmonic cues alone (Sloboda & Gregory, 1980).

Still more important, given that in general a click-location task requires subjects to provide a response *after the processing has taken place*, the click migration behaviour that emerges from those studies may not be the result of factors that operate *during* the processing of the constituent structure, but instead could originate in events that occur *after* the perceptual analysis has taken place. The fact that the task requires the participant to listen to the stimulus in the first place and simultaneously operate with representations of linguistic materials (written sentences / musical scores) in order to produce the delayed location response should elicit potential sources of interference in the given task. For example, the delayed performance compromises memory. The effect of memory could be minimized by involving subjects in tasks that require auditory input alone and the production of the response as close as possible to the hearing of the stimulus. However, previous attempts to modify experimental conditions in order to overcome this problem resulted in the emergence of other sources of interference. Therefore, it seemed to be impossible to separate the memory effects from the perceptual process *per se* using the click-location methodology.

Finally, interpretations of results in the application of this methodology seem to underline that the nature of participants' migration behaviour would be more a consequence of the lack of security in their judgments than a derivation of their representation of the constituent hierarchy (see Chapter IV for a more detailed account of this experimental technique).

Given that in the case of musical studies the click-location methodology was used to test the listeners' sensitivity to *phrase-structures and harmonic structures* as constituent organizations, using impoverished musical samples with a dubious musical value and few structural cues to "orient" the listener's attentional process, it was decided to apply it to see

if the above-mentioned shortcomings might be overcome when a hypothesis of constituency of *prolongational structures* using real musical fragments was proposed.

V.4.1.1 Hypothesis

In experiment V.4 a click-location paradigm was used in order to test the following hypothesis:

If a prolongational unit tends to be perceived as a coherent unit, it will tend to resist interruptions. Thus, clicks located at prolongational boundaries will not be experienced as migrating to other loci, but clicks located before and after prolongational boundaries will tend to migrate toward the prolongational boundary.

V.4.2 Method

V.4.2.1 Subjects

31 professional musicians, 16 males and 15 females, volunteered to participate in the experiment. Their average age was 28 years and their average musical experience was of 14 years. Musicians were graduate professionals with in degree of music at the university or at the conservatoire with at least 10 years of instruction on a musical instrument and a mean number of 6 years of instruction in music theory.

V.4.2.2 Stimuli

The same 11 musical fragments used in experiment V.2 were employed in this experiment. The treatment of the stimulus was identical to the one followed in experiment V.2 except for the location of the clicks.

The prolongation boundary was defined as the focal point between the last note of the prolongation and the following note in the progression.

In each musical fragment three click-positions were established:

Click-position at the prolongation boundary: click located in the middle of the inter-onset interval between the last note of the prolongation and the following note of the progression.

Click-position before the prolongation boundary: click located between the penultimate and the last note of the prolongational unit.

Click-position after the prolongation boundary: click located between the first note and the second note of the following unit.

The mean inter-onset interval between clicks in the three click-positions was 365ms.

As in experiment V.2, a hypothesis of relative salience of the beats constituting the metrical structure was also considered and the metrical position of the last note of the prolongational unit was manipulated.

V.4.2.3 Apparatus

Specifications concerning apparatus were similar to those of experiment V.2.

V.4.2.4 Procedure

The experiment began with a warm-up session, in which participants received instructions about how to proceed with the required activity, including rehearsal of the click-location task. In each trial, they listened to the melodic fragment three times in the following sequence: 1) fragment presentation without click; 2) fragment presentation with click, and 3) fragment presentation without click. During the first stimulus presentation subjects were required to listen carefully to the musical fragment focusing their attention on its pitch events' organization, that is to say, on aspects of melody, harmony, tonality, etc; during the

second sample presentation they had to pay attention to the location of the incoming click, and, finally, on the third presentation they were asked, as they listened to the fragment, to produce a click-location response (pressing a key on the keyboard) in the temporal locus where they believed they heard the click appearing during the second stimulus presentation. They were requested to provide the answer as precisely as possible, in order to avoid click-location errors and key-pressing errors. Once the warm-up session was completed, the test was administered to participants in two listening sessions with a ten minute interval between. At the end of the experimental session, they filled in a form about musical experience and were also required to write down the procedures that allowed them to solve the task during the test session. Subjects were tested individually.

Given the concerns mentioned before about the numerous sources of interference that could affect the click-location activity, the following precautions were taken for the experimental procedure:

i)   Non-inclusion of a familiarization session, developing instead a three fragment exposure presentation in which the subject was required to concentrate on those aspects of tonal pitch organization of the musical fragment that may orient the click-location response

ii)   Involvement of professional musicians only as experimental subjects, to ensure more accurate and quick structural processing

iii)   Presentation to subjects of an exclusively aural experimental condition, that involved them in the production of the location response without distracting their attention with the use of potential sources of interference

iv)   Avoidance of written support (musical scores) as a potential source of interference, given that the visual analysis of the written notation might lead to a biased click-location response, possibly activated by the use of musical knowledge that interfere with the feature under investigation. For example, the visual identification of the boundaries of *formal-phrase* units might collapse with subjects' awareness of the *prolongational* units which cognitive processing was the target of the current study

v)   Non-inclusion of a recognition task, given that subjects needed to attend thoroughly to each musical fragment with the aim of providing the click-location response; therefore, distracting their attention with the performance of a secondary task that required the use of additional cognitive resources would have interfered with the primary task.

V.4.2.5 Design

The test employed 44 trials organized as follows:

-11 trials with click position at the prolongation boundary and weak-strong metrical position at the prolongation boundary;

-11 trials with click position at the prolongation boundary and strong-weak metrical position at the prolongation boundary;

-11 trials with click position before the prolongation boundary;

-11 trials with click position after the prolongation boundary.

In study 3, therefore, four experimental conditions were set up:

a. Click-position at boundary

b. Click-position before boundary

c. Click-position after boundary and

d. Click-position at boundary, changing the metrical position of the notes at the boundary. These resulted in d.1 click-position at the boundary in weak-strong metrical position; and d.2 click-position at the boundary in strong-weak metrical position.

V.4.3 Results

Data were analyzed as follows:

The inter-onset interval between the two notes that surrounded the click locus was measured in milliseconds. Subjects' answers were assessed as to whether they fell inside or outside the inter-onset interval, and the answers (in milliseconds) were converted into categorical responses according to whether they fell inside or outside the interval.

Three possible answers were identified for each of the four conditions: 1: before the interval; 2: at the interval and 3: after the interval.

Although it was expected that click-positions *c* and *d* would generate the greatest number of migrant location responses *outside* the click-position interval, and that click-positions *a* and *b* would generate the greatest number of responses *at* the click-position interval, differences between the frequencies of clicks located to 'after', 'before' and 'at' the boundaries were not significant.

V.4.4 Discussion

The results do not seem to support the migration effect in the experience of prolongational structures. It was not possible to identify a clear migratory tendency in any of the four conditions, *except for answers in condition c* (click-position after the boundary). In this condition it is possible to observe something of a tendency to migrate backwards to within

the prolongational unit. This tendency may be informative about the different structural statuses of the two notes at the focal point of the "prolongation-progression" locus, reinforcing the evidence provided by results of experiment V.2 about the cognitive status of the prolongational boundary.

In experiment V.4, clicks were located around the boundary area in three positions: before, at and after the boundary. Following previous experiments in music perception that employed the click-location technique it was decided to use an in-between-notes criterion instead of a fixed temporal span criterion between click positions. Consequently, clicks were located in the inter-onset interval *between the tones* before, at, and after the prolongational boundary. Clicks at *before-boundary* position were placed between the penultimate and the last note of the prolongational unit; clicks *at-boundary* locations were placed between the last note and the first note of the following unit and clicks at *after-boundary* locations were placed between the first and the second note of the subsequent constituent unit. As noted before (see stimulus section), the mean of the temporal span between clicks was 365 ms. Given that perception organizes attention at intervals, a possible interpretation of the lack of differences might be that this interval magnitude had not resulted as long enough as to represent "a consistent temporal-span percept" between two different positions at the boundary region of the structural unit (Snyder, 2000).

On the other hand, almost all the subjects explained in their final written comments that in the procedure they followed to produce the location response they used a rhythmic-metrical strategy (counting beats, counting the number of formal units, assimilating the click into the rhythmic pattern in which the signal was superimposed, etc.) in order to undertake the click-location task, in spite of the experimental requirement of concentrating their efforts on the pitch organization of the musical piece. It seems as if the click-location

task may have elicited a rhythmic-metric surface representation instead of an underlying melodic-harmonic one.

A distracting task was not included in experiment V.4 because it was considered that it might interfere in the process of memorising the musical material. However, in spite of the precautions undertaken to set experimental conditions that did not provide a number of interference sources in the production of the click-location response, it seems that this experimental technique appears to be most sensitive to a kind of implicit information-processing in which the subject co-opts into a specific puzzle-solving strategy. Given that in the click location study, the musical samples that were tested were the same as those used on the click detection study, and, nevertheless, different results were obtained, these outcomes suggest that it is the click location technique which appears as problematic, rather than other factors. Similar shortcomings were found when this technique was used in previous investigation (see Chapter IV).

Thus, it is possible to assert that the paradigm of click-location seems not to be appropriate to study prolongational aspects of tonal music, at least in the experimental contexts in which it has been used so far. If the aim was to investigate further experimental conditions that fit better to this paradigm it might be interesting to try musical stimuli containing less evident metrical cues, in order to see if it helps orienting participants' focus towards the linear aspects of tonal music.

V.5 Conclusions: summary of the experimental findings and future prospects

In these experiments the constituent status of prolongational structures was investigated, and experimental evidence of listeners' experience of them during music attending was

found. Experiment V.1 found evidence of the cognitive status of constituent boundaries when sensitivity to closure was investigated. This evidence allowed the proposal of an hypothesis of constituency to prolongational structure; accordingly, its cognitive status was investigated in experiments V.2, V.3 and V.4. In the four experiments reported here participants were involved in real time music attending activities listening to fragments of real music.

Sensitivity to prolongation was tested by applying two experimental techniques that are more commonly used in experimental work in the linguistic field: click-detection and click-location. Clicks were located in the area of the prolongational boundary of structural units. In experiments V.2 and V.3, behavioural evidence of the cognitive status of prolongational structures was collected using a click-detection task. In experiment V.4, on the other hand, the use of a click-location technique did not exhibit a clear migratory effect in click-location responses, except for a small tendency to migrate clicks backwards to the boundary region of the prolongational units.

In what follows, these findings will be discussed in relation to the analysis of prolongational structure in tonal music, according to classic views in music cognition. The following topics will be discussed:

i) the status of constituency of the prolongational structure and the identification of the prolongational boundary;

ii) the problem of the experimental paradigm and the conception of temporality in music;

iii) the search for evidence.

V.5.1 The status of constituency of the prolongational structure and the focal point of the prolongational boundary

The experiments reported in this chapter studied the constituent status of prolongational structures at surface reduction levels. Sensitivity to prolongation was tested around the boundary area of prolongational units, and produced evidence of the cognitive status of this focal point during music attending.

It is important to highlight the fact that in this experimental section the main interest was in the relationship between constituency and prolongation. This is the reason why sensitivity to the boundary region was tested in two different contextual scenarios: that of embedded and linear constituent organizations (experiment V.1) and that of prolongational organizations (experiments V.2-V.4). In order to demonstrate that prolongational structures functioned as constituent organizations, it was necessary first to identify a clear constituent feature, to investigate its cognitive status, and once evidence was found, to locate this same feature in a prolongational scenario, observing if participants' behaviour was similar to the one demonstrated in the previous experimental situation. The structural feature under investigation was the boundary area. If a prolongational unit functions effectively as a constituent, testing listeners' attention to the area of the prolongational boundary would reveal something about its cognitive experience. The experimental work of study V.1 made possible to assert that effectively the boundary area *is* a focal point with cognitive value when ongoing listening takes place.

Prolongational theory understands structural units as entities in which prolonging pitch events cohere strongly on the basis of their dependency on the structural tone that governs them (see chapter IV for a detailed explanation of this concept). An hypothesis of constituency can be attributed to prolongational structures as long as they are capable of

forming structural entities in which pitch events play specific roles. For example, one of them should be able to operate as a structural tone (the head of the constituent) on which the rest of the pitch events are dependent all along the structural unit, until a new structural tone is reached. But in doing so, specific conditions need to be set up in order to assign the continuity of the unfolding voice-leading a capacity to activate a constituent parsing mode. For example, prolongational units must be conceived as having beginnings and endings. However, in theoretical accounts of prolongation, descriptions of the continuous linearity that result from the unfolding voice-leading do not clearly imply that boundary areas are delimited in a way that is similar to the way constituent theory suggests. Thus, for example, it is not expected that in prolongational structures constituent boundaries are necessarily delimited by the same structural markers that are in play in formal phrase units (see above a detail in the introduction of this chapter). In experiment V.2, the metrical component of the prolongational boundary was manipulated, changing the metrical position of the boundary note: nevertheless, sensitivity to that focal point was strongly found in both conditions, meaning that structural priority in prolongational structures rests more on the non-temporal information of the pitch component organization than on the rhythmic organisation of the time-spans of tones.

V.5.2 The problem of the experimental paradigm and the conception of temporality in music

Overall, reductional theories that emerged from cognitive science and linguistics understand reductional representation as the result of processes deployed by an 'ideal' competent listener. However, few specifics are provided about the way those processes are instantiated as long as music unfolds. In the experiments reported here the emphasis was placed on testing reductional representation in real-time listening contexts. The experience

of unfolding music on time is intrinsically related to the way attentional processes work. Music unfolding prompts expectations about incoming information, and attention is therefore related to the way temporal intervals are filled in, while listening to the musical piece progresses.

Two on-line experimental techniques, click-detection and click-location, were employed to test participants' sensitivity to prolongational structures. Click-detection proved useful in fulfilling this goal; on-line detection of extraneous signals located in the constituent unit and immediate production of mechanical responses seemed not to interrupt the ongoing listening process, simultaneously providing indirect evidence of the cognitive status of different events in the musical organization.

Click-location, on the contrary, was not useful in accounting for listeners' sensitivity at the boundary location, except for a weak tendency to migrate clicks backwards from the after-boundary position to the boundary one. The results confirm previous concerns about the shortcomings of this experimental technique in this context (see section IV.3.3. in Chapter IV). In spite of the various controls that were developed in order to alleviate the effect of these shortcomings, the experimental situation seemed to elicit in participants the deployment of a different listening strategy than that expected using this experimental procedure. In experiment V.4, subjects clearly developed a rhythmic-metric strategy when they were required to locate the click aurally. This behaviour elicits a rather different interpretation of results compared to previous explanations of those obtained when this technique was applied. It tackles an interesting issue in relation to the analysis of the temporal experience of music.

As was noted before (see Chapter IV), music attending seems to be related to time estimation, that is, to the way temporal intervals are phenomenologically 'filled-in'.

However, the sense of time in music seems to be dual: on the one hand, it is related to the chronometric experience of time; on the other, it seems to be integral to the intrinsic configuration of a particular piece of music. The former concerns a clock-like experience of time and implies the division of the musical continuity into regular beats, while the latter emerges from the individual's internal experience of the piece as a whole (Epstein, 1995). Both temporal dimensions shape musical time and are assumed to be available when music is experienced. It is probable that the click-location technique elicited in participants a clocklike experience when they had to locate a single 'atom' (the click) within the continuous unfolding of musical structure.

On the other hand, it is at the core of prolongational theory that the manifestation of the underlying structure unfolds temporality in a narrative sense (see chapter I for a detailed explanation). This conception of the underlying unfolding is related to the second dimension of the sense of time. Therefore, it is possible to think, according to results of experiments V.2-V.3 and V.4, that the click-detection task does not interfere with the on-line sense of temporal continuity that is more congruent with the experience of the unfolding voice-leading; as long as it takes place, the boundary of a prolongational unit appears naturally to the perceiver as a focal point at which a relative sense of completion is reached. SRTs accounted for such structural feature. On the other hand, click-location activity requires the participant to search for the occurrence of an 'atomic' event in the ongoing continuity of music unfolding. This demand elicits the emergence of the clocklike sense of musical time, filling-in the temporal span of the constituent unit with the concurrent strategies of counting beats, counting measures, assimilating clicks to the rhythmic structure of the piece, etc. In the end, the apparent lack of accuracy identified in previous studies appears as the result of the elicitation of a temporal dimension that is not

appropriate for the explanation of temporality in the sense that is proposed by prolongational theory.

V.5.3 Final remarks

The first section of this thesis (Experiments V.1 - V.4) has explored the cognitive reality of prolongational structure. Behavioural evidence of its experience was produced. Constituent boundaries were investigated in different types of syntactical contexts in order to look for evidence of their cognitive statuses as focal points in real-time music-attending situations. A click-detection procedure was useful in suggesting the cognitive reality of the focal point boundary of prolongation-progression in musical fragments that presented surface-reduction closed prolongations. Differences between musicians and non-musicians were not found in their sensitivity to prolongational structures, confirming previous findings about commonalities in music cognition when listeners with different degrees of musical expertise are investigated at a basic level of cognitive activity.

As was underlined in Chapter II, a hierarchical model gains credibility and status if its predictions have been rigorously tested, and if experimental evidence can plausibly account for the explanations it provides. Moreover, given that in the artistic experience, a given phenomenon sometimes elicits alternative interpretations, the plausibility of eliciting multiple responses during the experience of a given musical event is at the core of the nature of music. Therefore, the indirect evidence for internal processes and representations collected by means of experimental research is strengthened when converging results are obtained from different sources of experimental work. In these experiments, sensitivity to prolongational structure has been tested in adult populations of musicians and non-musicians. Also, replications of the experiments with children found ontogenetic support

for this cognitive process. All this converging evidence presents a strong case for the cognitive status of the prolongational structure in tonal music.

However, the current findings are constrained by the reductional level at which the experience of prolongational structure has been investigated. Even though the experimental techniques employed were useful in accounting for the on-line attentional behaviour of participants, they are not appropriate if the aim is to explore the cognitive awareness to deeper levels of the underlying structure. The shortcomings of this experimental method were demonstrated when a click-location paradigm was used. In that task, participants had to produce a delayed focal response that nevertheless involved them in an aural analysis of the global unfolding of the underlying structure for as long as a complete phrase. This result warns us about the efficacy of moment-to-moment experimental methodology in accounting for the understanding of more extended, global, processes of music unfolding.

On the other hand, prolongational structure in music is generally held to arise through the experience of patterns of tension and resolution (Lerdahl &Jackendoff, 1983, p. 179) that are related to the ebb and flow of the underlying structure in tonal musical pieces. It has been proposed that such patterns of tension and resolution afford one possible pathway to the experience of meaning in music (Steinbeis & Koelsch, 2007), one that is bound to, and is predicated solely on, the contrapunctal-harmonic structure of the music. However, it can also be proposed that the patterns of tension and resolution that have been identified as operational in musical perceptions arise by virtue of mappings between such patterns and other domains of human cognition. Larson (1997) has suggested that aspects of the perception of the ways in which melodic structures unfold in time can be accounted for in terms of the experience of physical forces such as gravity, inertia and magnetism. In a later paper, (Larson, 2004), he compares the results of experiments on melodic perception with

those of a computational model that addresses melodic 'movement' and expectancy in terms of such forces. Lakoff (1987) puts forward a theory of experiential meaning, suggesting that many 'primitives' in human cognition can best be understood by reference to body-image schemas deriving from the generalised constraints on our embodied experience of the world.

The ontology of prolongation - as it was proposed for the first time (Schenker 1935-1979) - is intrinsically related to aspects other than a strict constrainment to the acoustic component of musical compositions. The concept of prolongation is germane, for example, to philosophical foundations that understand the unfolding of music as being governed by some biological principles emerging from the natural world. The metaphor of organicism is present in the language used by Schenker to describe the composing-out of the underlying organization of musical pieces. Traversal of a path, directed motion, different forces in operation, achievement of a goal, are some of the metaphorical concepts found in Schenkerian writings to account for this idea.

This aspect of musical deployment – in which the musicological conception of the underlying structure is intrinsically embedded - will be interrogated in the following chapter, in order to pose a hypothesis of the experience of prolongational structure as a *structural metaphor*. Accordingly, we propose to explore the issue of whether the experience of patterns of tension and resolution that emerge from the composing-out of prolongational structures might be underpinned by such general body-image schemas; were this to be the case, it could provide a means of understanding the experience of music in ways that would enable relationships between musical experience and broader aspects of cognition to be clarified.

CHAPTER VI

EXPERIENCING PROLONGATION AS A STRUCTURAL METAPHOR

The mind is inherently embodied
(G. Lakoff and M. Johnson, 1999)

VI.1 Experiential realism: An alternative view of reality and meaning

VI.1.1 Externalism and internalism: Two arguments in the analysis of reality and meaning

In Chapter II, some issues relative to the status of perceived reality and evidence in the study of music experience were discussed. As noted there, when the status of reality of the human experience is considered, the arguments exposed suggest different answers to questions concerning i) the building up of categories about reality, ii) the ways in which the mind accesses knowledge, iii) the extent and types of mental processes that are taken into account and iv) the relationship between knowledge and the body. Their consideration is important to the purposes of this thesis because concepts in music theory, such as fundamental entities, like that which is the object of investigation here, or other general concepts, such as the concept of tension and relaxation, tend to be very vaguely defined; and while the previous chapters have demonstrated unequivocally that processes such as prolongation are cognitively relevant, they have not clarified sufficiently the epistemological bases of such processes in cognition. Clarification of this issue requires

considering two philosophical perspectives that have exerted and still exert influence on the study of the nature of human experience and knowledge. The contrasting arguments provided by them answer differently the problems involved when the status of reality is considered, as mentioned above.

On the one hand, human experience may be informed by the philosophy of *objectivist realism*. This externalist approach of reality postulates that the relevant factors of truth are different from the internal states of the individual. They are the result of a strict correspondence - in one and only one sense - between the symbols of thought (words and mental representations) and the way that the world *is*: the objectivist model of cognition understands the mind as a mirror of nature. As a consequence, the categorization of objects is seen as being *independent* of the mechanisms of human perception and cognition, excluding experiential aspects of human nature. The objectivist paradigm had an enormous impact in most of the contemporary philosophical theories of mind and language (e.g. the model-theoretic semantics by D. Lewis, 1972 and the semantic theory of truth and meaning by D. Davidson, 1975; 2001). Given that objectivist's assumptions have been taken for granted, the entities and symbols that are their representatives are assumed to be natural.

However, when attempts are made to define the theoretical concepts of music in any way, this is often in metaphorical terms (cf. Lerdahl and Jackendoff's (1983) account of tension and relase in terms of the ballistics of a ball thrown up into the air, or Schenker's conception of the underlying voice-leading as the traversal of a path). It is a strong assumption in this thesis that this metaphorical component in the theoretical explanations about prolongation can be investigated in terms of its capacity to convey meaning of the cognitive experience of the prolongational structure.

The philosophy of objectivist cognition, as it proposes an externalist view of reality, can not properly answer this issue. There is an alternative, defensible philosophical position, the philosophy of *experiential realism*, which might provide a better means of addressing issues such as the basic level of music categories and the role of imagination in the building of essential knowledge in terms of the 'metaphorical' processes that appear to underlie many areas of human cognition.

According to the philosophy of experiential realism (Lakoff, [1987]-1990; Putnam, 1981; Putnam, 1990), human cognition is informed by an internalist view of reality, in which the states of the mind are manifest through processes of perception and thinking that seem to provide a frame to build coherent meaning of the world. The internalist approach posits that there are aspects of cognition that may not have a relation at all with truth, meaning, reason or knowledge in the terms posited by the externalist, objectivist view. According to Lakoff ([1987]-1990), the building of mental categories is constrained by the fact that, out of the properties that objects possess, some are essential, that is, they are properties that make things to be as they are, and others are accidental, in other words, they are properties that things appear to have, but that are not the properties that capture their fundamental nature. Besides, there are categories that involve imaginative structures of understanding, such as schemata, metaphor and mental imagery. Furthermore, there are certain categories that have well defined boundaries and others that have fuzzy boundaries. Such is the case with categories that show prototype effects and idealizations. There are also categories that correspond to phenomena that do not exist in the external world, but instead are internal to us. That is the case of the category of colours, in which colours arise from our interaction with the world, but they are not outside us. Similar considerations apply to the categorization of sounds.

Those properties that, according to the experientialist approach, characterize concepts as belonging to the basic level (among others, Gestalt perception, whole-part organization, motor-movement activity, and image formation) could not truly be properties in an objectivistic theory, given that they depend on our experience with the environment. Contrary to the objectivist view, according to which basic categories correspond to primitive concepts or building blocks, without any internal structure, and complex categories result from the combination of such conceptual primitives, in the internalist view, those categories that correspond to basic concepts have, in many cases, an amount of internal structure of a kind that humans find it easy to process (to learn, to remember and to use).

The experientialist approach states that we understand reality in coherent ways that are founded in our embodied experience of the environment. Our conceptual system is based first on those concepts that are experienced directly; that is to say, on simple spatial concepts such as, for example, the concept of *up*. The structure of our spatial concepts arises from the recurrent nature of our spatial experience. Consequently, the ways that the mind is used to understand the world are, in part, based on the practical knowledge derived from basic conceptual structures that are built by means of intense environmental interaction.

Music, as a human activity, is also the result of the interactive communication between individuals and their cultural environments. The nature of musical experience, the ways in which musical meaning is built, and the kind of mental representations that are formed as a consequence of the practice and exposure to music, can be subjected to the same kind of investigation as other areas of human experience and meaning. Some research has already been conducted from this perspective (Larson, 1997, 2004; Saslow, 1997-1998;

Zbikowsky, 2002, Johnson and Larson, 2003, among others) although it has mostly been accomplished in theoretical terms.

In order to proceed to the analysis of embodied aspects of cognition applied to the problem of the experience of prolongational structure in music, there is a need to consider first some issues relative to the building of musical meaning. This will serve to situate the inquiry about the status of the prolongational structure in a wider context of meaning attribution.

VI.2 The nature of musical meaning

VI.2.1 Referentiality and musical meaning

According to some of the prevailing Western views about musical meaning, music seems to be self-referential. "As music unfolds in time, it articulates complex structures that relate to, and perhaps refer to, each other" (Cross 2005, p.32). Abstraction of its intrinsic meaning will thus depend on the extent to which the listener is able to realize inferences about music unfolding, based on what has already been and is actually experienced and, on the basis of the former, to create expectations about what is going to occur. This is what Meyer has called the 'evident meaning' of music (Meyer, 1956). Those evident meanings that account for music's auto-referentiality are linked to its intrinsic structure, and are unveiled as long as music unfolds in time. Although it is possible to hypothesize that there are similarities in the ways intrinsic musical meanings are conveyed in different cultures, and that those commonalities would be based on universals of the human condition, it is also true that the kinds of musical practices that have taken place in different epochs, or that occur inside different cultural environments, are shaped by their particular characteristics, and consequently it is expected that those cultural differences elicit the attribution of different meanings.

"The inferences that are made in the abstraction of evident meaning from music appear to depend on individual and cultural histories, and on both generic and specific attributes of the cognitive systems that make them" (Cross 2005, p. 33).

On the other hand, music not only seems to refer to itself, that is to say, to what emerges from the inside of its inherent structure, but also to convey meaning that would be the result of its location in the context of a wider framework of human experience. For example, the experience of the unfolding of musical structure can be attributed to the capacity to modulate our emotional responses to music (see Cross, 2005; Juslin and Sloboda, 2001; Jackendoff and Lerdhal, 2004; Mauss, 1988). It is probable, according to Meyer (1956), that music *has* a connotative referential capacity to elicit the emotional domain. This referential capacity is assumed to be linked to metaphorical interpretations of human understanding and may constitute an analytical tool to account for analogy in music (Cumming, 1991).

Moreover, according to Cross (2005), music has 'floating intentionality', that is to say, it possesses a capacity to bear different meanings to and between performers and listeners. Music's floating intentionality becomes a fertile milieu in which a process of cross-domain mapping can be instantiated. A metaphorical interpretation of meaning effectively occurs by means of the establishment of correlations between different experiential domains (see below for a description of the concept of cross-domain mapping). It is assumed that this propensity of music is a function of human development that is based in the proto-musical behaviors that allow the re-description of information between different domains (Karmiloff-Smith, 1992) and that emerges as a consequence of the first experiences of intersubjective parental-young infant interaction (see Lakoff and Johnson, 1999;

Trevarthen, 1999/2000; Stern, 1985; Meltzoff, 1981; Malloch, 2002; Pöppel and Wittmann, 1999; Dissanayake, 2000; Miall and Dissanayake, 2003).

In summary, the meaning embodied in music can be understood, on the one hand, in terms of its capacity to refer to itself, that is, its auto-referentiality (after Cross, 2005). This evident meaning that emerges from the inherent musical structure becomes real for the individual as long as music unfolds in time. On the other hand, music can be understood in terms of its referential capacity to bear metaphorical connotations that emerge from the peculiarities of human experience, thanks to their capacity to reflect the temporal forms of the brain-mind-body processes (see Damasio, 1994; 1999). Those general attributes that are inherent to music, such as its sonic organization, quality of movement, heterogeneity of meaning, basis in social interaction, and personal meaning, elicit a definition of its nature as a domain that "embodies, entrains, and transposably intentionalizes time in sound and action" (Cross 2003, p. 24). Even in the context of passive exposure to music, it seems that the activity of listening involves the activation of brain regions that are linked to the domain of the body movement (Janata and Grafton, 2003, in Cross, 2005).

At a basic level of experience, musical meaning does not appear to be grasped so very different from the ways in which other types of basic categories are built. Facilitated by the social and cognitive flexibility that characterizes humans (Cross, 2005; Karmiloff-Smith, 1992) some imaginative structures are developed and used to build musical meaning. The metaphorical proclivity of music to express concepts or images that are deeply linked to the individual's embodied experience of the world, and the potential links between human imagination and the understanding of the prolongational structure, are the focus of the present chapter. The way metaphorical thinking is developed in the context of human experience will be treated in the following sections.

VI.3 Human cognition at the basic level

VI.3.1 Cognitive categories and experiential knowledge

Understanding reality requires, in part, being able to think in terms of categories of events. However, as complex as the world can be, the ways reason and inference are used to structure understanding seem to be based, as we saw above, on a level of categorization that is neither the lowest nor the highest, but is located in the middle of taxonomical descriptions, called the 'basic level' (Rosh, 1973; 1978; Rosh et al, 1976).

Basic level categories are developed in the course of our environmental activity. Investigation of the basic level claims that this cognitive dimension is intrinsically dependent on the way spatial configurations of the physical world are perceived (Lakoff, [1987] - 1990). From the point of view of a theory of embodied knowledge, categories at the basic level represent hierarchical systems neurally realized, endorsing cognition with an inferential and / or imaginative capacity to bound pieces of information coming from different experiential domains. Basic concepts are structures that we use to form our essential categories. "An embodied concept is a neural structure that is actually part of, or makes use of, the sensorimotor system of our brains. Much of conceptual inference is, therefore, sensorimotor inference" (Lakoff and Johnson 1999, p. 20). It is not yet clear whether conceptual inference uses the same brain structures as perceptual motor inference. It seems that our inferential capacity is evolutionarily adapted, in order to enable to fit our bodies to certain extremely important differences in the natural environment. By means of embodied action, a kind of enactive cognition is developed (Varela, Thompson and Rosch, 1993) and the basic-level categories are built as a consequence of this process. Knowledge at the basic level is, therefore, the consequence of some peculiarities of the human design,

including aspects of our bodies, brains and minds, such as mental images, gestalt

perception, motor programs and knowledge structure (*cf.* Lakoff and Johnson, 1999).

The basic level equates to the cognitive level at which the members of a given category

share forms and aspects that are similarly perceived; it is the level at which a single mental

image can represent a whole category, and it is also the level at which most of our

knowledge is organized (Rosch, 1973; Rosch *et al*, 1976; Mervis and Rosch, 1981, Gopnik

and Meltzoff, 1992; Johnson and Mervis, 1997) At the basic level, people name things

with simple labels, remember them easily and perceive them holistically. For example the

category 'chair' is in the middle of the categories 'furniture' and 'rocking-chair'.

According to the characterization of the basic-level, mid-level categories are cognitively

"basic" (Lakoff and Johnson, 1999, p.27). However, the basic level is variable (Rosh,

1978). It is dependent on the cultural context and the level of expertise, and it varies

according to whether categories are 'natural', in the sense that they emerge as a

consequence of the environmental interaction (Tversky and Hemenway, 1984) or

'artificial', meaning thay they are purposefully created to perform particular types of

analyses in the laboratory. In the last case, they are not typically an analogue representation

of the natural world; instead, membership is based on the stipulation of necessary and

sufficient conditions of this type: the members of category x must have features y and z,

and if something has features y and z, then it is a member of category x. However,

according to Zbikowsky (2002), artificial categories simply represent specialized forms of

natural categories. For example, we find a connection between the natural category *group*

that emerges when the sequence of sounds of a musical piece is segmented into significant

units, and the artificial taxonomies established by formal music theory.

As said at the beginning of this section, categories at the basic level correspond to cognitive constructs of a relatively higher order and are the result of myriad other processes that operate at lower levels of human consciousness. They are essential to the processes of reason and inference. The group of categories that form the basic level accounts, not for a one-to-one correspondence between 'atomic' elements, but between groups of attributes that are understood as unified wholes, as gestalts. Grasping meaning at the basic level depends partly on the feasibility of the body to mediate between the properties of categories. The types of categories that the basic level deals with are not only basic categories about objects but, as was said before, they also refer to other basic constructs of human experience, such as mental images, motor programs, social concepts and basic emotions.

Processing at the basic level could explain, in an appropriate psychological way, the cognitive salience of musical features, such as for example, the salience of the musical motive as the minimal comprehensive unit of a musical piece, given that in this musical event the focus is on the whole more than in its atomic particles (Zbikowsky, 2002). It is an assumption of this thesis that the underlying structure of a piece of music, understood as a one-to-many correspondence between music events, can be assigned a status in music cognition in the terms proposed by psychological descriptions of the basic level.

As long as human environmental interaction takes place, it seems as if knowledge at the basic level is organized, in part, around cognitive constructs called prototypes (Rosh and Mervis, 1975). They seem to govern the operation of imaginative or inferential tasks that involve the use of specific categories. A prototype contains the features that are statistically more relevant (in cognitive terms) amongst the members of a determined category, against which other potential members are compared. A basic prototypical category is a relational

network formed by a number of attributes that are weighted according to their typicality to define membership within the category. Prototype effects are the result of comparisons between the attributes of the individuals to be categorized and those of the exemplars stored in long-term memory (see Chapter III for previous research in the study of the underlying structure as a prototypical category).

A standard analysis of prototypicality, considers that the attributes that are more salient are the attributes that are more recurrent. However, some empirical work on categorization has shown that the standard analysis of prototype effects exhibits certain shortcomings that question its full validity to account for the process of categorization at the basic level (Barsalou, 1985; 1989). For example, Zbikowsky (2002) performs an analysis of prototype effects over the first thirty-seven measures of the fifth symphony from Beethoven. He compares the initial motive and the successive varied presentations of such musical material through the whole passage. Once the analysis is performed he finds that it partially fits what he assumes to be the impression that a Western competent listener might have stored in her/his memory of Beethoven's motive: in such a stored impression, what counts as important as *the* prototypical memory register is the first presentation of the motive with its characteristic gesture in tutti and fortissimo. Less important are the subsequent presentations of the same motive with variants in the magnitude of those attributes. According to the result of this analysis, the assumed prototypical expectations were only satisfied with respect to the melodic contour, but do not match the other attributes of dynamics and instrumental density. Therefore, it is argued that we are able to keep the initial motive as the mental image that synthesises the prototypical gesture of the piece, in spite of the number of instances in which such a motive is presented with other orchestration and dynamics during the first thirty-seven measures of the movement. The

difference between the analytical evidence and the assumed memory register is due in part to the constraints imposed by the method followed in the analysis of prototypicality: it takes into account the frequency of occurrence of the attributes of events as the determinant of the degree of prototypicality in a given category. The salience of an exemplar is analysed performing statistical countings of what is more *common*, that is to say, of the recurrent values assigned to the attributes of the individual members that are being considered. But such frequencies of occurrence might not represent the values that are considered by the individual listeners as more *important*.

It follows from the above that some attributes are more important than others to define the structure of a given category. For this reason, then, if we want to focus on the attribution of meaning to the musical experience, the issues pertaining to the prototypical nature of knowledge, considered just in terms of frequency of occurrence, are a necessary but not a sufficient condition to account for the ways in which meaningful structures are formed in music. There are *other* structures that operate at the basic level where categories are built, and they are also useful to perform inferences and to reason about the components of reality. Those structures are imaginative in nature and will be described in the following section.

VI.4 The embodied nature of human cognition

In the previous section it was proposed that categorization at a basic level is one of the most pervasive forms of meaning attribution to our experience in the world. It was also noted that prototypicality is a primary factor on which to build categories at the basic level. Finally, it was posited that some imaginative structures that exist at that cognitive level

would also operate in the building-up of categories about reality. These structures will be discussed in what follows.

According to Lakoff and Johnson (1999) concepts, categories and internal experience are bound together to account for the imaginative and embodied nature of human cognition. Human cognition does not consist uniquely in algorithmic relationships between abstract symbols and objects in the world in the terms posited by the objectivist view; it is also the result of the use of imaginative and embodied structures that emerge from our sensory-motor experience and guides the development of conceptualization, reason and inference (Lakoff, 1994).

Traditionally, embodied aspects of human experience have been ignored as a source of understanding. This is, in part, due to the argument of the externalist view that perception and movement, on the one hand, and concept formation and reasoning, on the other, are completely separate faculties. But experimental evidence from cognitive linguistics (Gibbs, 1994; Johnson and Lakoff, 2002; Kemper, 1989; Gibbs and O'Brien, 1990; Gibbs et al, 1994) suggests that some imaginative structures, developed from our bodily experience, are used in our understanding, inference and reasoning processes. According to the evidence collected, they *are* a part of our conceptualization system. Another source of evidence that throws light on the plausibility of the embodied-mind hypothesis, meaning that perception and movement *are* a part of reasoning, comes from simulations in the field of neural modelling. Results show that perceptual and motor mechanisms take part in the performance of abstract inferences (Lakoff and Johnson, 1999; Feldman and Narayanan, 2004; Narayanan, Chang, Feldman and Narayanan, 2004). What those simulations show is the plausibility of the existence of forms of coordination between brain activity and determined corporal states through a global neural mapping; this parallel stimulation of

different groups of neurons in different areas of the brain may account for the shared capacity of motor and conceptual structures.

According to the above, some abstract components of thinking should be modelled by certain human physiological characteristics, such as being vertically oriented or sensing the action of determined forces when movement is experienced in environmental activity. As long as we receive stimulation from the external environment, the so-mentioned global mapping should create a dynamic circuit in the brain that continuously relates our gestures and postures with independent sets of sensory signals. Also known as the dynamic core hypothesis (Edelman, 1992; Edelman and Tononi, 2000) this notion emphasises that consciousness is both integrated and differentiated by a cluster of ongoing complex activity at the neuronal level. It is by means of this process of continuous interpretation of bodily information that the mind may build the basic cognitive structures that will then be used to develop knowledge about reality. Johnson (1987) has called this capacity *humanness*, referring with this term to the human ability to apprehend reality that is the foundation of the imaginative structures embodied in our mind.

VI.4.1 Image - schemas and metaphorical projections

According to the embodied knowledge view, the basic units of perception should not be single and isolated occurrences of components, but of groups of events that unfold dynamically in time. This kind of Gestalt perception is assumed to be the foundation of our everyday-life basic knowledge and involves the operation of some cognitive structures called image-schemas (Johnson, 1987; Gibbs Jr. and O'Brien, 1990; Turner, 1993; Gibbs Jr. and Colston, 1995; Kreitzer, 1997; Lakoff and Johnson, 1999).

Image-schemas are recurrent structures that operate in our perceptual system. They originate as the result of the spatial and temporal experience of our bodies in movement. They develop as a consequence of the physical manipulation of objects in the environment. That is to say, image-schemas are based in our direct, kinaesthetic activity. But they have an internal structure and an internal logic. They function somewhat like the abstract structure of an image, and thereby connect up a vast range of different experiences that manifest the same recurring structure. However, image-schemas are not rich, detailed images. Furthermore, they are by no means visual. The idea of an image is simply a way of capturing the recurrent organization inferred from the patterns of behaviour in the process of concept formation. For example, the *verticality* schema (Johnson, 1987) is the abstract structure of those recurrent experiences in which an up-down orientation is involved.

Image-schemas are characterized as dynamic analogical representations of spatial relations and of spatial movements. Although they derive from motor and perceptual processes it is argued that they are not themselves sensory-motor processes (Johnson, 1987; Lakoff and Johnson, 1999). Neither are they passive containers in which experience is kept. Rather, they are resources by means of which we build order in our experience. Therefore, image-schemas are rather different from the notion of schema used traditionally in cognitive science (cf. Rumelhart and Ortony, 1977), where they are considered as abstract and propositional conceptual structures (see Chapter IV). Nevertheless, they share with traditional schemas the characteristic feature of describing molar aspects of the psychological domain, that is, they refer to processes of a relatively higher order of neural organization, in opposition to those molecular aspects that focus on the activity of isolated neurons (Thomas, 1999). Image-schemas are imaginative and non-propositional in nature, and operate as structures that organize experience at the level of embodied perception and

movement. Image-schemas are not necessarily static, as in the case of the above-noted traditional definition coming from cognitive science. That is, they are not empty templates that are filled-in with sensory input. As Neisser suggests, they have a moderate degree of flexibility, that is, they are malleable structures of perception and motor programs (Johnson 1987, p. 29). According to Johnson this definition shares certain commonalities with Kant's idea of schemas as structures of imagination that connect concepts with percepts (see Chapter I for preliminary accounts of Kant's conception). In words of Reybrouck (2001) "[imaginative projections] hold a position between the epistemological paradigms of *realism* and *nominalism,* that stress either the sensory 'realia' or the imaginative reconstructions of these realia in the listener's mind" (p. 117). Image-schemas are present in different sensorial modalities, such as visual, aural, tactile and kinaesthetic. They are not isolated entities, but are frequently related to each other to perform certain operations of transformation that play a significant role in the relationship between perception and reasoning.

In summary, image-schemas are recurrent structures formed by elements that are related according to an internal logic that makes them to appear as kinds of gestalts. They operate in the perceptual system and are built as a result of the environmental interactions and motor programs. They provide structure and coherence to the spatial and temporal experience of the body in movement manipulating objects in the environment. Image-schemas are the product of the kinaesthetic experience. They function in a similar way to the abstract structure of an image, but they do not form rich and detailed images. They organize mental representations, connecting a vast range of different experiences that exhibit the same recurrent structure.

Each image-schema has a very simple and clear internal structure, suitable to be applied to a wide variety of instances of our everyday activity. We can see a representation of the most pervasive image-schemas in Figure VI.1.
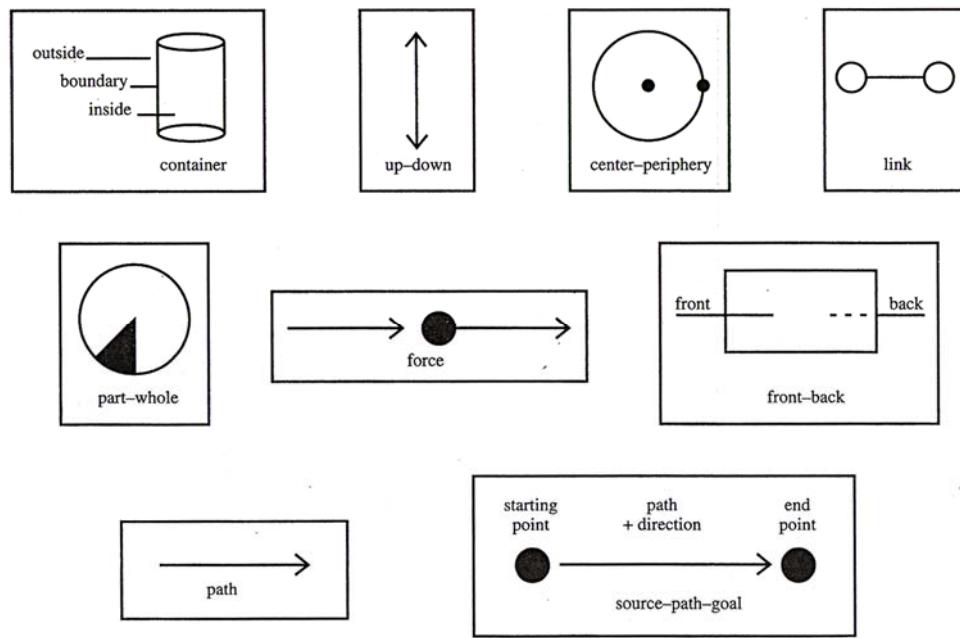
Figure VI.1 Diagrams of the most common image-schemas.Extracted from Saslow (1997-1998, p. 219.)

A relatively small group of primitive image-schemas are believed to be responsible for providing structure for those concepts that involve spatial-relations in language use. Among others, we find up-down, in-out, part-whole, centre-periphery, link, cycle and near-far image-schemas. Spatial-relational concepts also contribute to structure dynamic components in the group of force schemas, such as pushing, pulling, propelling, support and balance. Components of spatial orientation in spatial-relation concepts include vertical, horizontal and front-back image-schemas (Lakoff and Johnson, 1999).

We apprehend spatial relations performing bodily projections. We use movements of the whole body and of the body members in relation to the surrounding environment. Through

the operation of our sensory capacities, a set of basic spatial orientations are defined and used to perceive the relationships between different objects. In the interaction with the surrounding space we project, for example, the idea of our body as a *container*, onto other containers, either closed, such as a room, or open, such as a garden, and we say that we are *in* the room or *in* the garden. These linguistic expressions are forms of embodied cognition that are the result of our bodily interaction with the environment, and are referred to as phenomenological embodiment (Lakoff and Johnson 1999, p. 36). There is another type of embodiment called neural embodiment that describes the neural mechanisms that give rise to concepts, such as for example, the structure of colour categories.

The theory of image-schemas is crucial for the explanation of the phenomenon of *metaphorical projections*. In metaphorical projections, kinesthetic image-schemas that represent meaningful structures of direct environmental experience are mapped onto other more abstract domains that are not experienced directly, with the aim of structuring them (Lakoff and Johnson, 1999). For example, in the case of the image-schema *container* a metaphorical projection occurs in the following statement: '*I don't want to leave any relevant data out of my argument*'. Here *argument* is an abstract entity that is metaphorically grasped as a container, which, in turn, is partially structured by the *out* schema. The primary metaphor involved here is *argument is a container,* one of the most common ways of understanding the concept of argument in Western culture (Johnson 1987, p. 35). Thus, metaphorical projections are processes of transformation by means of which image-schemas evolve into extended structures of organized meaning at more abstract levels of cognition, such as inferential reasoning.

Our general conceptual system appears to be metaphorical in nature. According to this view, metaphors should not be understood exclusively as literary resources, but also as

agents that operate in our conceptual organization (Lakoff and Johnson, [1980] - 2003). Metaphors are indeed *metaphorical concepts*. This is the reason to define a structural metaphor as a *cross-domain mapping* in the conceptual system (Lakoff, 1993, p. 203). In summary, structural metaphors are among the most prominent cognitive structures by means of which we make sense of our experience in the world. They are related types of embodied, imaginative structures that are used to project patterns from one domain of experience to another domain of a different kind, with the aim of structuring it. They are vehicles used to organize our understanding, projecting from the concrete to the abstract the ordered patterns of spatial, temporal and physical domains of human experience.

Metaphorical mappings seem to appear as a consequence of an early ability to perceive abstract similarities between events that are neither physically similar nor associated through co-occurrence (Wagner et al, 1981; Phillips et al, 1990). According to Lakoff (1990) and Turner (1990; 1993), mappings between domains in metaphorical projections do not refer to the imposition of a structure from a source domain onto the target domain, but to the establishment of correspondences between them. Those correspondences are not haphazard; on the contrary, they preserve the image-schematic structure latent in each domain. It could be argued that those correspondences are due to the principle of invariance developed by Lakoff (1990). This states that metaphorical mappings of information between source and target domains preserve the structural characteristics of the cognitive topology of both domains. For example, in the metaphor *more is up* (Johnson, 1987), such propositional expression accounts for a number of experiential connections that are not themselves propositional in their origin. In *more is up* quantity is understood in terms of the *vertical* schema. Thus, *more* and *up* are correlated in our experience in such a way that provide physical bases to our abstract understanding of

quantity. Thus, a cross-domain mapping process is observed when the *up-down* schema is mapped onto the *emotional* domain: here the image-schematic structure of the physical states we associate with emotions is preserved, and the spatial orientation of the underlying *verticality* schema is imported in order to provide a coherent account of those states. The cross-domain mapping process is reflected in some of the expressions we use when we refer to those emotional states. For example when we say: *"I'm feeling up today"*, there is a correlation between the affective state of being happy and the energetics coming from the upright posture. The conceptual metaphor that underlies that surface statement is *happy is up* (Lakoff and Johnson 1999, p.50). The phrase suggests not a literal representation of the spatial domain, but a use of our knowledge of the physical space to structure the understanding of emotions.

A distinction should be made, then, between conceptual and linguistic metaphors. A conceptual metaphor is a cognitive mapping between different domains; a linguistic metaphor, instead, is a linguistic expression of such cross-domain mapping. This significant difference between both kinds of constructs is important to understand the use of metaphorical thinking in music; it will help differentiate between a metaphor as a cognitive process of attribution of meaning by means of a mapping between the physical domain and the musical domain, and the linguistic expression of such mapping, that is, the use of metaphorical language in the theoretical discourse about music.

An unanswered anthropological question is whether culture precedes or is a consequence of the operation of image-schematic structures and metaphorical projections in the formation of concepts (Saslow, 1996). Lakoff (1990) seems to assign structural metaphors a fundamental role in the construction of thinking. Metaphorical thinking also guides the development of empirical theorizing and scientific research (Fernandez-Duque and

Johnson (1999). According to this view, metaphorical mappings are natural in the sense that they are motivated by the structure of our experience. In order to contribute to the development of the anthropological argument, it can be useful to distinguish between image-schemas and metaphors as occurring at different levels of our conceptual structure. Image-schemas and their bodily-derived structures are very basic, even though their peculiar manifestations in language and conceptual structure might vary among different cultures. This statement makes sense if it is assumed that image-schemas are driven from our shared embodied experience of spatial orientation. However, in order to determine the functioning of metaphorical projections through different languages and cultural experiences, it is necessary to be sensitive to the conceptual systems that underlie them (Pederson et al, 1998). For this reason, an analysis of imaginative thinking in the context of our musical cultures must take into account the conceptual models that are predominant in Western theoretical discourses about music.

In summary, according to the view that is proposed here, embodied cognition originates, in part, from the development of some preconceptual and imaginative structures, called image-schemas, that emerge from our kinaesthetic experience and that are considered essential in the development of our making sense of reality. Since environmental interaction involves the development of an increasingly rich network of image-schematic metaphorical projections, human experience can be understood as a wide domain that includes a number of cognitive dimensions that organize our construction of meaning and reasoning. Musical concepts, as a subset of abstract concepts, can also be understood as derivations of the recurrent patterns of spatial orientation of our bodily movements; such patterns would be put into action when we experience music in different types of activities, that is, during listening, performance, composition or thinking about music. This

interwoven correlation between sound perception and movement will be focused on the following sections.

VI.5 Music cognition as an embodied process

VI.5.1 Cross-domain mapping as a metaphorical process in theoretical discourse about music

Metaphor is present in the language used to describe music (see Zbikowsky, 1997, 1997-1998, 1998, 2002; Saslaw, 1996, 1997-1998; Guck, 1991,1994; Larson, 1997, 2004; Brower, 2000; O'Donnell, 1999; Spitzer, 2004). As noted above, those metaphorical linguistic utterances used in music are surface realizations of conceptual metaphors that are basic structures of knowledge, built in the mind by means of mapping processes that establish correspondences between different experiential domains.

The use of metaphorical thinking can be observed, for example, in the linguistic descriptions of the temporal nature of music. As was highlighted before in this chapter, it is argued that musical meaning is inherently linked to the temporal unfolding of the sonic stream of events of a musical piece (Lochhead, 1989/90). According to Johnson and Larson (2003), music 'moves'. And this metaphorical description of the temporal aspects of music seems to be expedited by its link with the experience of movement in other domains of our life. The meaning embedded in musical motion connects the sense of musical time with the manifestation of time and movement in different types of activities. The acquisition of the capacity to listen and to respond properly to musical temporality seems to be based on such correspondences.

The understanding of musical motion is particularly affected by the fact that our concept of time seems to be structured by, or is at least inextricably linked to, our physical experience of space (Johnson and Larson, 2003). For example, those linguistic expressions used to talk and think about music as moving from-and-to different musical locations seem to find their origin in that assumption. We conceptualize time as spatially structured, and ascribe attributes of movement in space to it; thus, temporal intervals become distances, and specific locations in the musical piece become goals that are reached; in other words, they are treated in terms of destinies at which we arrive. We encounter obstacles in our journeys, or we find ourselves blocked with respect to our final destination, or we need to apply different kinds of forces in order to initiate and sustain motion. For example, let us consider one of the musical dimensions involved in the analysis of musical motion: pitch. Concerning *pitch movement*, for example, it is frequent to find descriptions of the fundamental relationship between the tonic and the dominant tones in terms of *movement from* the tonic *to* the dominant and then of *return* from the dominant to the tonic; also, that the tonic is *reached* or is *arrived at* (see Saslaw, 1996).

It has been proposed that our understanding of the temporal unfolding of music can be conceptualized in terms of three metaphorical dimensions: *music in movement, musical landscape* and *forces in movement* (Johnson and Larson, 2003). According to the first of them, a musical event is conceptualized as an object that moves in time and space with respect to a stationary observer, in this case, the listener. Describing music in terms of the *moving music* metaphor, means to think our experience of a musical piece as sharing something with our experience of seeing objects (musical objects) move in physical space. The second dimension refers to a basic experience of physical motion that refers to our ability to move our bodies through a spatial landscape. In this experience, the musical work

is conceived as an extended three-dimensional landscape through which the listener moves. Musical events are the locations on such musical landscape. The third metaphorical dimension involves the experience of music as physical motion; it is related to the idea of being moved by physical entities from one point to another. In music, the metaphorical force is the music itself; acting on listeners to move them from one state-location to another, along some path. Thus, one can actually feel as being '*pushed', 'pulled'* and generally '*moved'* by the music. The *music as moving force* metaphor helps to explain why we use the word 'move' to mean these different, but related, things.

These metaphorical dimensions seem to be based in our dynamic experience of movement and physical forces, and also seem to account for the way in which the logic of physical movement corresponds to the logic of musical movement. It is suggested that these metaphors constrain the inferences that are realized when musical motion is reasoned; that is, they define *what* is that moves, the *way* it does, and the *location* towards the movement takes place.

To the extent that our understanding of time is also deeply metaphorical in nature, conceptualization of the temporal dimension adds another component to the analysis of musical motion. In Western culture, we conceptualize the experience of time as movement in space. Research in cognitive linguistics shows that time is conceptualized in terms of two basic spatial metaphorical systems: (i) objects that move themselves towards a stationary observer; this happens when we say, for example, *Christmas is coming*, and (ii) the observer moves her / himself all along a landscape or scene, and times are now the locations or regions in such scenario; in this case we would say, for example, *we are approaching the deadline.* Both metaphorical systems define most of our building of time spatialization. They constitute reversed examples of figure-background: in the former, the

times are the figures that move in relation to the static observer (as background) while in the latter the observer is the figure that moves in relation to the temporal scene as background. Both metaphors are based in a conception of temporal unfolding as relative spatial movement, and they also play a central role in our understanding of musical motion.

We experience physical movement in space on three principal ways: i) we see objects moving; ii) we move our body and iii) we sense our body moved by forces. These pre-conceptual experiences elicit the development of a wide body of knowledge about the physical attributes of movement.

Given that musical motion, like physical motion, occurs in time, the basic metaphors about musical motion include metaphorical conceptualizations about time, space and physical features. Statements such as *the reprise is coming* or *the strings go faster now*, are metaphorical ways of describing the movement of musical events towards us or passing in front of us. A future musical event is something that is coming; once it has occurred it exists only in memory, in a metaphorical space behind us. Therefore, the metaphorical experience of music's temporal unfolding combines the notion of physical trajectories of movement with that of times in movement. Statements such as *we are arriving at the coda*; or *we are now in the second movement*; or, *the flute solo is waiting to enter in measure seven*, exemplify the process of cross-domain mapping relative mainly to the metaphor *musical landscape* that conveys the above described experiential dimensions.

Relevant to the analysis of the experience of time in music is the view that Imberty (1981, 1987;1991; 1992a,b; 1993) develops about the temporal organization of the musical piece. He posits that a musical composition exhibits a dynamic quality that determines our experience of its temporal unfolding in terms of indices of relative stability. Assuming that the musical form is experienced as a schema of its temporal structure, that is to say, as the

reduction of the continuities and discontinuities of tensions and distensions in the sequence of events of the musical piece, or, moreover, as a mental representation of its temporal progression, he tests this hypothesis by comparing the listeners' experience of the temporality in musical pieces of different composers such as Brahms, Debussy and Berio (Imberty, 1981;1987; 1990). He finds that there are two types of experiences of temporality: (i) a concrete, vivid experience, where time is processed in the instant-to-instant sequence of events, of sequence of contrasts and of progressional passages, or of other continuities of any type in which events are ordered in an irreversible manner, and (ii) an abstract experience of time-space, in which the formal and logical relations of the musical piece are intertwined in such a way that allow the listener to go back and forth, thanks to the expectations of the incoming unfolding of events or the recapitulation of what has been previously experienced. Finally, Imberty posits that in tonal music those dynamic aspects of the temporal progression that are tied to the narrative thematic organization of the composition elicit a more abstract experience of time-space, in which a temporal schema of relation of order is activated, while in non-tonal music, the absence of thematicity and the relative salience of the continuous changes of the event's sequence are determinant factors to prompt a concrete instant-to-instant experience of time, in which a schema of order is activated.

Saslow (1996; 1997-1998) explores the role that the physical experience of forces plays in shaping some of the music-theoretical formulations which appeared in Schenker's treatises. Applying the theory of image-schemas and metaphorical projections just described, she portrays a theoretical description of the way some of the most salient metaphors of the Schenkerian approach are structured in a conceptual system derived from movement, spatial orientation and physical relations that imply force.

The *force* schema is discussed in detail by Johnson (1987). Among the primary features of the force schema are the following: force is experienced through interaction; force usually involves the movement of some object through space in some direction; there is typically a path of motion involved; forces have origins or sources and they can be directed to targets, they have degrees of intensity or power, and finally there is a sequence of causality involved (Johnson 1987, p. 43-44). The *force* schema develops as a consequence of our causal interaction with the environment. Environmental activity requires the exertion of force as we act upon objects or as we are acted upon by them. Image-structures of force come to play a central role in our capacity to understand the physical experience. Seven category members of the force schema have been hypothesised. They are: *compulsion* (exerted force with a given direction, a magnitude and moving along a path); *blockage* (in the attempt to interact forcefully with objects, to encounter obstacles that block or resist our force); *counterforce* (two equally strong force centers collide face-to-face with the result that neither can go anywhere); *diversion* (a force vector is diverted as a result of the causal interaction of two or more vectors); *removal of restraint* (removal of a barrier or the absence of some potential restraint); *enablement* (a felt sense of power to perform an action) and *attraction* (a kind of gravitational force towards an object). In Figure VI.2 the seven members of the force image-schema are shown.
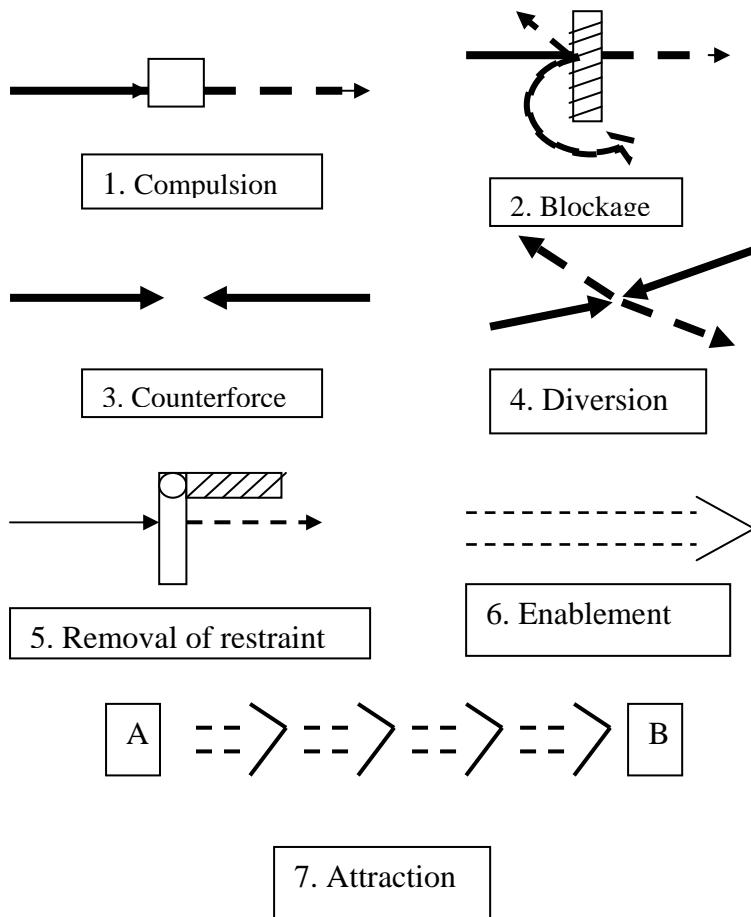
Figure VI.2. Diagram with the seven members of the force image-schema:1. compulsion; 2.blockage; 3. counterforce; 4. diversion; 5. removal of restraint; 6. enablement, and 7. attraction (adapted from Johnson, 1987).

Saslow understands force as a driving image-schematic structure that conveys meaning to descriptions of the underlying musical structure. She analyses the theoretical ideas developed by Schenker in his work *Free Composition* ([1935]-1979) with the aim of identifying the subtypes of force image-schemas that underlie Schenker's utterances. She finds that Schenker's statements display many of the features of the *force prototype* and there is also evidence of the use of some of the members of the force category, among them *compulsion* and *blockage* (see also Lakoff, 1993) in his analytical descriptions. It is observed how the embodied physical experience of movement and orientation constrains

253

Schenker's conception of the underlying structure as a live force. Schenker's graphic representations of a musical piece were also studied by Saslow. She found that his graphic system also depicts some of the image-schemas examined earlier.

In summary, Saslow's theoretical analyses indicate, as was posited in Chapter I, that Schenker's metaphorical conceptualizations of the underlying hierarchy in tonal music might be the surface manifestations of an idealistic conception that could be explained in the terms posited by the theory of image-schemas and metaphorical projections.

The analysis of the nature of musical knowledge as an embodied process highlights the fact that there is not a single ontological perspective with which to approach musical experience. On the one hand, there are commonalities of embodied cognition that account for the presence of universal features of perception and movement that help build musical knowledge. However, the analysis of the embodied experience is also informed by conceptual differences that can be identified in the musical cultures throughout different epochs and places. The phenomenon of conceptual change accounts for a rather complex analysis of *the musical experiences* that emerges from different cultural environments. This also underlines the idea that musical structures do not exist in the objective world independently of the person who experiences them; instead, it is the individual who makes a strong contribution to their existence by means of the processes that s/he deploys to convey meaning of her/his experience with music. These cognitive structures that are involved in knowledge-building organize themselves in terms of concepts and relationships between concepts, taking the form of conceptual models. They are constrained by cultural determinants and are used to make sense of reality. They can plausibly provide answers to the problem of the relationship between the universal and the particular in the construction of meaning. Coming back to the analysis of Schenker's metaphorical descriptions, it is

evident that they depict a conceptual model containing the idea of the underlying unity of a musical composition and its manifestation in the composing-out process of the voice-leading as unfolding.

Before going deeper into the analysis of the Schenkerian conception in the terms proposed by the philosophy of internal realism, the next section presents a description of the significance of conceptual models and its relation to the process of cross-domain mapping in music.

VI.5.2 Conceptual models: categorization in music

As was highlighted before, some aspects of music organization can be described by employing linguistic expressions that contain conceptual metaphors. The use of those linguistic resources makes apparent the existence of an underlying knowledge base that results from the combination of music cognition as embodied knowledge, and its projection onto the conceptual system used to characterize music.

When those conceptual descriptions are considered as a whole, it is noticeable that the linguistic discourse accounts for a cognitive organization that can be interpreted in terms of what is called a conceptual model (Zbikowsky, 2002). A conceptual model is formed by a group of concepts belonging to a specific knowledge domain that are linked by specific relationships. Conceptual models are stored in memory as a unit and recovered as wholes to respond to particular environmental cues or stereotype situations of reasoning. They are part of a variety of cognitive models, among them, mental models, cognitive idealized models, cultural models and mental spaces, that have been developed to account for those cognitive structures that are involved in our understanding of the world (Lakoff, 1990). The primary function of a conceptual model is to provide a guide for categorization and

reasoning. The simplified representation provided by a conceptual model selects groups of attributes according to their value of typicality in the context of the categorization task. For example, the conceptual model used to categorize the form of the fifth symphony (see above) reflects our intuitions that there exists only one primary motive in that musical composition. Therefore, the analysis of typicality would be based on the impression that the initial musical motive functions as an anchor in the process of elaboration of Beethoven's piece.

Global conceptual models develop as a result of the abstraction emerging from clusters of more local conceptual models. Global conceptual models are, in turn, embedded in the wider knowledge base that constitutes a given musical culture. In the case of the fifth symphony, the local conceptual model about the musical motive is informed by those features of a more global model that constitute thematicity in music, and the more general global model that corresponds to Beethoven's particular musical theme reflects the influence of the German and Austrian musical culture of nineteenth century (Zbikowsky, 2002).

The conceptualization of pitch is another interesting example of the use of conceptual models that depends on cultural constraints. Pitch is a cultural concept characterized differently according to place and time: for example, in Western tradition it is metaphorized as high and low, reflecting the mapping of the up-down spatial orientation. But in other cultural environments it is characterized differently: for example, in Java and Bali it is understood as little and large, by analogy with the size and velocity of the sound-source vibration; among the Suya from Amazonas the equivalents of high and low are young and old, based on the belief that the vocal register goes down with age (Zbikowsky, 2002). In all these examples a mapping between domains takes place, encouraging some

conceptualizations and dismissing others. Coming back to the experience of pitch in the Western tradition, for example, the cross-domain mapping becomes plausible because the image-schematic structure of spatial orientation is projected onto components of the acoustic domain. Thus, *more is up* correlates with the structure of pitch frequency. Both domains are continuous and can be divided into discrete elements (points in the spatial domain and pitches with determined frequencies in the acoustic domain). The vertical-space representation of musical pitch creates a conceptual abstract space, a kind of imaginative metageometry that functions as a relational system informed by associations between the geometric possibilities of other cognitive spaces that provide knowledge of localization, direction, distance and reference system (Duchez, 1979). In this learned system representation, the linguistic expressions of high and low are the surface manifestations of a metaphorical phenomenal reality that is founded on basic correlations between the above-mentioned sensory domains.

In spite of the described commonalities of our primary embodied experiences which have been described, the variety of conceptual metaphors used to characterize our understanding in a given knowledge domain leads to the problem of why a given mapping is used more than another. The reason has precisely to do with the conceptual models that we apprehend in the context of our own culture. The Western conceptualization of pitch as high and low goes back as far as when polyphonic music was written for the first time (Zbikowsky, 2002). It is also related to a global model that considers pitches as objects, that uses graphic symbols to visualize them, and that preserves the musical pieces in the format of the Cartesian conception of the relationship between time and space (Douchez, 1979). In other cultures there are other constraints that characterize pitch conceptualization; for example in Java and Bali there is no difference between a sound source and a pitch

(Zbikowsky, ibid). The apparent dissociation between pitch and sound source in the West does not, however, eliminate the image-schematic basis of their understanding. It only suggests that a cultural process of pitch reification developed retrospectively in the context of the Western musical literacy. It is significant for the analysis of the conceptual models developed inside a specific cultural environment that the intense academic debate that frequently takes place in the core of a given domain, in our case, the musical domain, is more about the ways those categories that account for the organization of musical comprehension are defined than about an absolute definition of music.

> "The specific mapping chosen within a tradition of discourse about music reflects not so much absolute musical structure as it does the broader cultural practice within which music and its understanding are embedded: mappings reflect the conceptual models that are important to culture. The cross-domain mappings employed by any theory of music are thus more than simple curiosities - they are actually key to understanding music as a rich cultural product that both constructs and is constructed by cultural experience" (Zbikowsky 2002, p. 72).

In synthesis, cross-domain mapping processes are not accidental or arbitrary; their development reflects two conditionings: (i) the conformance to the invariance principle by means of which the mapping preserves the image-schematic structure of source and target domains (Turner, 1990; Lakoff, 1993) and (ii) the idea that the mappings are constrained by the conceptual models that frame them. For these two reasons, the analysis of the cross-domain mapping process is essential for our understanding of the theoretical discourse about music. For example, the different theoretical perspectives of the concept of tonality as developed by music theoreticians such as Riemann, Schoenberg, Kurth and Schenker, that reflect logic and the psychological principles of energetics, symmetry, gravity, etc., represent *ways of understanding* the musical organization (see Zbikowsky 2002, p76.). They reveal that musical understanding is deeply modelled by the metaphorical and

conceptual frameworks that derive from the logic of our embodied and spatial experience. We cannot separate this embodied activity from our cognitive experience of music. We do not experience a musical piece and then understand it. Our understanding *is* the way we attribute meaning to our experience.

VI.5.3 Cross-domain mapping in music listening: the experience of the underlying musical structure as a metaphorical process.

In this section it will be hypothesized that listening to a musical piece is also an experience of a metaphorical nature. In this metaphorical process we use our embodied knowledge of temporal, spatial and physical relations and project it onto the musical domain in order to grasp musical meaning. In particular, the focus will be on the way we experience the underlying structure of music.

Recently, a theory of musical forces was developed, and its principles were applied to the analysis of tonal compositions (Larson 1997a, b; 2004). This theory, based (i) on Gestalt principles, (ii) on aesthetic postulates about the dynamic quality of the artistic expression (Arnheim, 1984), (iii) on the theory of metaphor above described (Johnson, 1987) and (iv) on some of the principles of Schenkerian theory, develops psychological and musicological arguments about the imaginative nature of music listening representation. An essential aspect of the theory relies on the way in which music unfolds over time, conveying a sense of direction that helps in structuring a tonal composition. The theory conceives music perception as an imaginative and/or creative process, by means of which it is possible *to hear x as y*. The listener creates meaning because, consciously or unconsciously, s/he assigns sounds to categories. This process is the result of the interaction between the structural quality that arises from the internal make up of the piece of music and the

listener's mental processes that s/he uses to assign meaning to it. Thus, the experience of a musical piece is understood as a dynamic process that originates in the energetic quality that emerges from the musical piece, which is mapped onto the dynamic patterns that emerge from the listener's embodied experience. The dynamic quality of this process of meaning attribution is the result of the interaction of three musical forces: (i) *gravity*, the tendency of an unstable tone to go downward; (ii) *magnetism*, the tendency of an unstable tone to move towards the closest tone - which increases as soon as the discourse reaches the final goal - and (iii) *inertia,* the tendency of a pattern of movement to continue in the same way, depending on how it is imaginatively listened. Paraphrasing Salzer ([1962] - 1982). Larson applies the term *structural hearing* to the process in which listeners attribute meaning to sound relationships under the form: "*to x hear as y"*; for example, to hear a pitch pattern *as* an ascending gesture. The underlying structure is a function of the process of *contextual stability*, defined as the interplay between the three musical forces cited above, as opposed to the *inherent stability* of the musical piece, which is derived from the principle of consonance. The relative structural weight of each of the tones of the musical surface depends on the contextual stability of that tone: to hear a tone as unstable is to *envision* (expect) a more stable tone the unstable tone tends to, and to imagine the path it follows to reach it (generally involving stepwise motion). Therefore, to hear a tone as unstable means to hear it as an embellishment of a more stable tone, *at a more remote level* of the underlying musical structure (see in Chapter I, a description of the concept of mental retention and in Chapter V the characterization of structural hearing as the capacity to interpret an event as related to a structurally more important event in the underlying hierarchy). Elements at deeper levels are "abstract representatives" of the occurrences at more superficial levels. The propensity to hear certain tone combinations as inherently

stable is an emergent property of the interaction of simple perceptual mechanisms, some of them being universal, and others culturally determined (Huron, 2001; see in Chapter II a more detailed description of these issues).

Lerdahl (2001) acknowledges the value of Larson's argument, as informative of the psychological plausibility of abstracting underlying features in the experience of tonal music. According to him, the musical forces proposed by Larson begin to build a psychological account of what Schenker called *the will* of the tones, that is to say, metaphorical intuitions relative to the capacity of tones to convey in the listener expectations of continuity and resolution of melodic and harmonic tensions. Although Larson's hypothesis has been applied to the analysis of melodic surfaces, that is, at the note-to-note level of the musical structure, the value of his theory of musical forces as applied to deeper levels of the underlying structrure remains more theoretical. For example, he identifies a series of tonal processes that occur at the note-to-note level and that are recurrent in tonal musical compositions (Larson, 1997-1998). The ways in which those melodic sequences unfold matches his theory of musical forces. More importantly, these processes represent a collection of prolongational patterns of the tonic note that can be found in hidden prolongations. As was described in Chapter I, a hidden prolongation, also called motivic parallelism (Burkhart, 1978) is the phenomenon in which a given melodic motive can be exposed at different structural levels. The analysis proposed by Larson is interesting to the extent that, on the one hand, it brings support to the idea of linearity as a building principle that penetrates the underlying hierarchy, and on the other, it accounts for issues of tonal coherence, to the extent that there are some melodic continuations that are more coherent than others. If the unfolding of such components obeys 'dynamic' constraints that are topologically similar to some of the constraints we

experience in the physical world, then a cross-domain mapping may very well be a plausible explanation of the metaphorical projections implied in the experience of the music unfolding. This analysis of tonal patterns and musical forces has been extended to other tonal alphabets, such as the dominant triad and the chromatic scale (Larson, 2004). This theoretical extension follows the cognitive-music-linguistic tradition that hypothesizes the representation of tonal structure in terms of alphabets (Deutsch and Feroe, 1981; Lerdhal, 2001; see a detailed explanation of this feature in Chapter II).

The metaphor of musical forces underlies the idea of music as purposeful movement. This idea of music moving with a purpose is reflected in linguistic descriptions such as 'D is a passing note between C and E', or 'large skips are balanced by step movements', while expressions such as 'that dissonance needs to resolve' or 'the climax is reached at G' emphasize the metaphorical idea of the will of the tones, adding further linguistic evidence to the metaphor of music pursuing a purpose. The important assumption here is that we do not only think *about* music in terms of structural metaphors when we talk about passing notes, willing of resolution, etc. but we *musically experience* the piece in terms of those metaphors. In the case of the metaphor of musical forces we map our experience of physical gestures onto musical gestures. Further linguistic expressions such as 'the melody *climbed* even higher' or 'the leading tone is *pushed* toward the tonic' reflect that gestural quality. Consequently, a fruitful avenue of investigation of the imaginative nature of the prolongational structure will be related to the formulation of hypotheses and the design of experimental conditions to test the activation of metaphorical projections at deeper levels of the underlying hierarchy. This is the goal that is pursued in the second experimental part of this thesis that will be presented in Chapter VII.

VI.6 Summary: the status of prolongation as imagined cognition

In this chapter it was proposed that imagination embraces human knowledge in different domains of experience, thereby providing structure for the understanding of reality. Structural metaphors, based on systematic correlations between experiential domains, organize and structure human knowledge, helping make sense of reality (Lakoff, 1993; Johnson. 1995; Gibbs Jr., 1996; Johnson, 1993; Johnson, 1989; Zbikowsky, 1997-1998). They possess properties that emerge from the prototypical nature of human experience. They are applied automatically and regularly to assess aspects of reality. They originate in the use of structures of properties that occur naturally in our experience as a consequence of the direct experience, including manipulation of the environment. It is by means of cross-domain mappings that metaphorical projections take place, in order to assign meaning to categories and concepts that we express, for example, in our everyday language (see Gibbs, 1994; Kemper, 1989; Lakoff and Johnson, 1999; Fauconnier, 1990; 1997). Their use is also reflected in the ways language is applied to conceptualize and theorize about music. It is also assumed that music cognition can be understood as an embodied experience and that metaphorical projections can occur during music listening (Lockhead, 1989/90; Zbikowsky, 1997; 1998) helping listeners to grasp the musical meaning (Gärdenfors, 1988; Brower, 2000).

Among the variety of different and possible meanings that are evident in a musical piece, there is one in particular that the literature on music theory refers to as tonal coherence. This is a kind of musical meaning derived from the unfolding of the tonal organization of pitch in time, and the structural relations between tonal pitches that emerge from such organization have been the foundations of the Western musical tradition for more than four centuries so far. The dynamic quality of music unfolding has been and is still understood in

263

some musicological sources in terms of what is called 'energetics' (Rothfarb, 2002), a property conveyed by musical motion that is analogous to the physical concepts of force and power. The origins of the metaphor of force in musicological sources can be traced at the beginning of the twentieth century and, since then, it has been used to characterize music, for example, as a thematic drama of forces, or as an interplay of kinetic energies, or as conveying a sense of musical spatiality, or finally as a property that is within the tones themselves. Consequently, the dynamic properties of the tones combined in the musical piece convey coherence to the musical form and, in the end, integrate its structure into a dynamic whole (Rothfarb 2002, p. 928). It is evident that the characterization of the structural force metaphor in the terms proposed by the theory of image-schemas and metaphorical projections is informed by this conception of musical force as energetics.

This analysis of structural metaphor as a factor that delineates human experience is instrumental in acknowledging a human conceptualizing capacity, understood as the ability to form symbolic structures (image-schematic and basic-level concepts) that correlate with preconceptual structures, to project those structures metaphorically from the physical, spatial and temporal domain onto structures in more abstract domains, to finally form complex concepts and general categories (Lakoff, [1987] - 1990, p.281). In other words, we are referring to a general capacity to form what is called *idealized cognitive models* (Lakoff, *ibid*). To propose a hypothesis of the metaphorical experience of prolongational structure is to understand that such experience is founded in (i): the generation of an idealized cognitive model, based on the image-schematic structures formed as the result of our embodied experience, and (ii) the metaphorical mappings performed between those structures and the peculiarities of music's intrinsic organization.

Idealized cognitive models (ICMs) are founded on the basic image-schematic structures that were described above. If, whilst listening to a musical piece, a person is required to perform a cognitive activity that, involving the use of categorical thinking, elicits the experience of the underlying musical structure, it is proposed that s/he will understand such organization in metaphorical terms to the extent that the operation of an ICM is activated in his/her mind. This cognitive organization will elicit the performance of a process of metaphorical projection in which the sonic sequence of events will be categorized as possessing some structural properties that correlate with those image-schematic structures that form the experiential basis of such embodied activity. In other words, if the listener is exposed to the experience of a prolongational organization, the peculiar characteristics of such structure will activate the mental configuration of an ICM that will allow *understanding the prolongational structure in terms of some kind of structural metaphor.*

As we saw before, although conceptual metaphors are abundant in theoretical writings about music, the use of image-schematic knowledge in the experience of listening to the underlying musical structure has not been tested experimentally so far (for preliminary studies on the use of metaphorical projections as analytical tools see, for example, Askness, 1997). In spite of all the above-mentioned theoretical and experimental arguments about the nature of the prolongational structure, the following questions remain: In what way is the prolongational structure hierarchical? How are underlying events abstracted? How does the listener derive those hierarchical structures? Here, the prolongational musical structure is hypothesised as an *imaginative construct* whose descriptive value needs to be investigated in order to actualise it as real or imagined sound.

In this thesis a hypothesis of the metaphorical experience of prolongational structure is proposed. The use of image-schematic abstractions and their correlation to structural

features of music, by means of a cross-domain mapping process, will be tested using as an experimental device a type of prolongational structure called the interrupted structure. The structural metaphors that are assumed to be operating in the experience of the interrupted structure are two of the subtypes of the force metaphor, blockage and removal of restraint. The results of the experimental testing are reported in Chapter VII.

CHAPTER VII

EXPERIMENTAL STUDIES PART2: CROSS-DOMAIN MAPPING AND THE EXPERIENCE OF PROLONGATIONAL STRUCTURE

VII.1 Introduction

In this chapter, it is hypothesized that underlying musical structure is metaphorically experienced on the basis of the implicit use of basic image-schematic structures developed during the course of bodily interaction with the environment. By means of a process of cross-domain mapping, the listener metaphorically projects knowledge from the embodied environmental activity to organize and structure information in the musical domain.

In the experiment that will be reported, the interrupted structure, one of the most pervasive types of prolongational structures in tonal music, was selected as object of experimental testing. A cognitive view of this structural principle is proposed in terms of the conceptual metaphor of musical forces. The hypothesis is proposed that the interrupted structure is metaphorically experienced in terms of the blockage-removal of restraint image-schema, a sub-schema of the force image-schema (Johnson, 1987). Once that image-schema was activated, by means of a cross domain mapping process, it projects itself onto the underlying organization of the tonal piece, priming a categorization process by means of which the listener *understands* the interrupted structure as a sonic unfolding of such force schema.

In order to arrive at the formulation of the experimental hypothesis for study VII.1, the main ideas developed in Chapters I and VI will be summarized below, and the experimental rationale will be developed.

VII.1.1 The status of the prolongational structure: an example of an idealized cognitive model?

The analytical tradition that flourished during the past century in the field of music theory developed a body of musicological works that proposed a description of tonal structure in terms of hierarchical organization. Some of them describe music hierarchy in a way that appears to be similar to some of the descriptions provided by models of human cognition (Cohen, 2000). They are called *models of underlying structure* or *reductional models*, in that they attempt to account for a level of experience at which a work is grasped as a single pattern or unitary structure, rather than as a concatenation of atomic elements, patterns or composite of parts (Benjamin, 1982). In these models, relationships between hierarchical levels take the form of a one-to-many correspondence.

Schenker ([1906] - 1990); ([1922] - 1987); ([1925] - 1994); ([1926] - 1996); ([1935] - 1979) was one of the pioneering musicologists who developed a model of underlying musical structure, understanding the hierarchy of a tonal piece as a one-to-many correspondence between tones. In his theory, the underlying organization is a harmonic-contrapuntal pattern that functions as an archetypal unity (see Schenker, 2004). It unfolds from the background to the foreground through prolongational procedures in the form of an inclusional hierarchy (Cohn & Dempster, 1992). The particular arrangement of pitch events that results from the different elaborative processes that compose-out the fundamental structure is called the underlying voice-leading. Prolongation is the result of

the operation of those linear processes, by means of which some structurally more important events govern other structurally less important events, in such a way that the latter extend the phenomenal existence of the former. Prolongation is characterized in Schenkerian sources (Salzer [1962] - 1982; Salzer & Schachter, 1969; Forte & Gilbert, 1982; Cadwallader & Gagné, 1998; Schachter, 1998) as a comprehensive expression that conveys different composing-out ideas, such as "the elaboration, development, manipulation and transformation of underlying principles" (Salzer & Schachter, 1969, p. xix). Implicit in the unfolding of the underlying structure is the idea that its temporal deployment is guided, according to Schenker, by an *organic force* that permanently strives to attain the tonal goals that govern the musical structure.

From a psychological point of view, the experience of the underlying structure is assumed to occur, in the first place, as a consequence of previously acquired stylistic information about the regularities of the tonal idiom that are stored in long-term memory. To the extent that this knowledge is available while listening to a tonal composition, it will implicitly prime the understanding of its unfolding in the terms proposed by reductional theories. However, the explanations developed so far about the cognitive status of the underlying structure are not sufficient to provide comprehensive answers to the issue of the ontology of prolongation as experienced. Given that, in general, the musicological descriptions of the structure of music are not devoted to explaining the ways in which people listen to music in everyday environmental contexts, but instead of how music 'ought' to be heard, that is to say, how musicians should develop the kind of sophisticated listening that is required as a competence in professional musicianship (Cook, 1989), the following questions remain: How exactly are the underlying events abstracted? How does the listener derive hierarchical structures from the musical piece? What is indeed a prolongational

269

representation in the terms proposed by the theories of underlying structure? How is organic coherence experienced during listening? The sections to follow will answer these questions from the point of view of the theory of embodied cognition.

VII.1.2 Embodied cognition

Recently, a number of hypotheses about the metaphorical nature of music cognition have been formulated, highlighting the idea that metaphorical thinking, present in the language used to conceptualize music, might to some extent, model musical experience. According to these (see eg. Lakoff & Johnson, 1999) imagination, traditionally related mainly to the domain of creativity, would seem to play a central role in the processes of understanding, reasoning and attribution of meaning which occur in our every day experience (Johnson, 1987). The image-schematic structures that organize our basic knowledge are acquired during the course of our sensory-motor activity, and are activated unconsciously when we are grasping meaning in different areas of knowledge.

It follows from this that imagination must also play a primary role in music cognition. Our understanding of music should also be in part a consequence of the embodied structures or forms of imagination that grow out of our bodily experience. Therefore, by means of imagination, we assign meaning to music while listening, performing, composing, or conceptualising it. In particular we are going to propose in this study that the way people understand and convey meaning in music is, at least in part, mediated by a process in which we use some basic image-schematic structures that have been developed through our environmental activity. By means of a metaphorical process, named cross-domain mapping (Lakoff, 1990; [1987] - 1990) we use knowledge coming from a given experiential domain, for example the physical domain, to understand information belonging to another

domain, in this case the domain of the sonic-musical structure. The experientialist approach should strongly support the assumption that the prolongational structure could be understood in terms of a cross-domain mapping process. Accordingly, if the listener experiences prolongation in metaphorical terms, h/she will assign categories to sounds, that is, h/she will understand pitch events sequences in terms of some of the image-schematic structures mentioned above.

Some of the theoretical postulates about underlying musical structure have already been proposed as conceptual metaphors (see for example Zbikowski, 2002; Saslaw, 1996; 1997-1998; Larson, 2004). However, the use of image-schematic knowledge in the listener's experience at deeper levels of the prolongational structure has not been tested experimentally to date.

In this chapter, the experience of underlying structure as a metaphorical abstraction will be investigated by running an experiment on categorization, in which it is expected that subjects perform a cognitive process whereby they understand the underlying structure in metaphorical terms (see Lakoff, 1987). Here we hypothesize a cognitive view of the voice-leading principles in terms of the conceptual metaphor of musical forces. This approach, founded on some of the ideas of the Schenkerian tradition, argues that the temporal unfolding of the underlying structure elicits a kind of imaginative listening by means of which a sense of direction is conveyed to the musical flux. Some of the linear techniques that unfold the underlying structure, producing the melodic and harmonic prolongations (see Cadwallader & Gagné, 1998; Salzer, [1962] - 1982), prompt the activation in music perception of an imaginative and/or creative process, by means of which it is possible to hear *x as y* (Larson, 1997) that is to say, to assign sounds to categories. This process is the

result of the interaction between some musical features that arise from the internal make-up of the piece of music and the mental processes activated by the listener.

In summary, the experience of a musical composition is understood as a dynamic process that lies behind the energetic quality that emerges from the sonic organization of the musical piece, which is mapped onto the dynamic patterns that emerge from the listener's environmental experience. The dynamic quality of the tonal composition arises from the interaction of tension-relaxation features of the pitch event's organization, for example, the tendency of the fundamental line to go downward, the tendency of an unstable tone to move towards the closest tone at different levels of the underlying hierarchy, and the general tendency of the musical discourse to reach the tonal goal. To hear a tone as unstable means also to hear it as an embellishment of a more stable tone, that is to say, to hear it as an embellishment of another tone at a more remote level of the musical structure (Larson, 1997). The tendency to hear certain tone combinations as inherently 'stable' is an emergent property of the interaction of simple perceptual mechanisms (Huron, 2001; Bregman, 1990): some of them being universals, and others culturally determined. According to Lerdahl (2001), the idea of musical forces in action begins to build a psychological account of what Schenker called the *will* of the tones, that is to say, intuitions relative to expectations and musical tensions.

In order to see if prolongational structure has cognitive reality, its cognitive status will be considered in the terms proposed by Lakoff and Johnson, when they define the concept of metaphorical thinking. Following this idea it will be proposed that structural metaphors, which result from the operation of a variety of image-schemas, organize music listening. Some musicologists (Saslow, 1997-1998; Rothfarb, 2002; Zbikowsky, 2002) found in Schenkerian theory linguistic examples of metaphorical thinking in the terms proposed by

Lakoff and Johnson. The aim is to see if those metaphors not only underlie theoretical explanations of the musical structure, but also organize the listening experience.

In particular, it will be posited that the interrupted structure, a Schenkerian prototypical construct derived from his conception of musical form, is understood metaphorically in terms of the blockage – removal of restraint image-schema. In the experiment to be reported, the relationship between the interrupted structure and this image-schema is investigated in the context of a music listening task, in which participants are primed with visual animations that represent different image-schematic structures. Such imaginative constructions are assumed to be active in perception during music listening. Once the image-schemas are activated, participants are asked to match musical fragments with different musical reductions of those fragments.

In the section that follows the concept of the interrupted structure is defined.

VII.2 The interrupted structure

The interrupted structure is a two-part organization of the musical phrase that establishes a kind of internal phrase division that takes the form of $\hat{3}\hat{2}//\hat{3}\hat{2}\hat{1}$ or $\hat{5}\hat{4}\hat{3}\hat{2}//\hat{5}\hat{4}\hat{3}\hat{2}\hat{1}$ in the fundamental line and I V//I V I in the fundamental harmonic progression (see Schenker [1935]1979; Cadwallader & Gagné, 1998; Forte & Gilbert, ([1982]-1992). The interrupted structure is a symmetrical antecedent-consequent phrase construction in which the antecedent phrase is articulated by a half cadence and the consequent phrase by its correspondent full cadence. The two parts of this structure illustrate a harmonic and melodic process that begins in the first part with $\hat{3}$ or $\hat{5}$ in the upper voice over I in the bass, and continues to $\hat{2}$ over V, where the tonal motion ceases. At this point the unfolding process is interrupted, articulating a division in the structure. Immediately thereafter, the

motion begins again and this time it achieves the goal, that is, $\hat{1}$ over I. This final resolution that was initially expected at the end of the antecedent phrase as a realization implied by the interruption, only takes place at the end of the consequent. It is important to notice that when the second phrase restates, the structure retraces its path from $\hat{3}$ or $\hat{5}$ over I, moving through $\hat{2}$ over V, and only when $\hat{1}$ over I (the goal of the *Urlinie* and the bass arpeggiation) is reached, the sense of completion and fulfilment finally ocurrs. In short, the second part of an interrupted structure resolves the tensions created by the interruption of the first part.

The process of interruption plays a significant role in shaping tonal structures over spans of music of different lengths. The essential melodic-harmonic path between I and V is usually fulfilled with other intermediate elaborations of the harmonic motion that typically occur. Schenker named the two parts of this tonal process as the two *branches* of the interruption. The goal of the first branch is V, supporting $\hat{2}$ in the upper voice. The dominant chord that serves as the unresolved goal of the structural motion is frequently referred to as a *divider* (*Teiler*) or *dividing dominant*. Because V contains the leading tone, that embodies a strong tendency to *ascend* to $\hat{1}$, and also $\hat{2}$ in the upper voice tends to *descend* to $\hat{1}$, it is easy to understand why so much tension is accumulated when the bass and melodic motions are interrupted at that point. When the second branch of the interruption begins, it retraces the path of the first motion to finally achieve closure and dissipate the tension of the incomplete first branch with the arrival to $\hat{1}$ over I. Sometimes the end of the first branch - the 'point' of interruption - can be extended; that is, $\hat{2}$ over V can be prolonged over larger spans of music. Such an extension can be used by composers as a means of creating form, for the prolongation itself may become a discrete section within the overall form. This procedure is seen especially in rounded binary and sonata forms.

It is not always easy to determine whether an interruption has occurred in a specific instance: not every half cadence, for example, necessarily indicates an interrupted structure. In many cases the second branch of the interruption, whether it be a phrase or a larger section, will begin identically or similarly to the first branch - both melodically and harmonically - creating the impression that the second branch 'answers' the first. In this sense there is a relationship between the two branches of the interruption and the conventional *parallel period* - an antecedent phrase followed by a thematically similar consequent (Cadwallader & Gagné, 1998). What is important to bear in mind when the interrupted form is analysed is *what it is* that is being interrupted: the descent (or 'journey') of the *Urlinie* on its way to $\hat{1}$. When $\hat{2}$ over V is reached in an interruption, all melodic and harmonic progressions cease, at least on the level at which the interruption occurs. This means that $\hat{2}$ at the end of the first branch is not a lower neighbour of $\hat{3}$ at the second branch, and V does not resolve to the immediately following I. If it did, there would be not a half cadence but an authentic cadence. There is, of course, considerable tension produced by an interruption, because, as was said before, $\hat{2}$ over V carries definite expectations. As was noted above, such tension is only resolved at the conclusion of the second branch, with $\hat{1}$ over I.

In summary, interruption is understood as an architectonic device (Salzer, [1962] - 1982) that helps build form at different levels of the underlying structure of a composition. By means of interruption, a structural unit is expressed in the form of two distinct parts. The standard unfolding presents a melody that proceeds to the tone above the tonic while the bass reaches V; at that point, both are interrupted in their movement to the tonic. The structure evolves from that point, generally presenting the same musical material until the final resolution is reached. That is to say, in order for a phrase to show an interrupted form,

it is necessary but not sufficient that its unfolding leads to the dominant; that is to say, not

any arrival at a dominant harmony is indicative of the presence of an interrupted structure:

in the interrupted form, the motion of the upper voice and the bass are interdependent.

Unity is guaranteed by the completion of the structure in the second phrase; in this
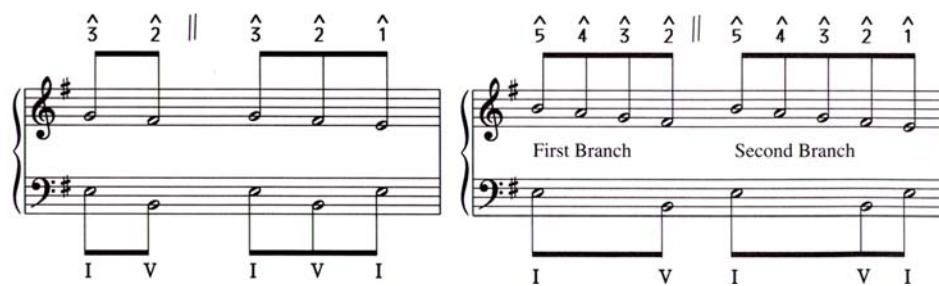
Figure VII.1 Schema of the two types of the interrupted structure: 3-2//3-2-1 and 5-4-3-2//5-4-3-2-1. (Extracted from Cadwallader and Gagné1998; pp. 166-167)

structure both parts form a whole at a higher structural level. Figure VII.1 shows a schema

of the interrrupted structure.

As stated before, the interrupted structure finds its musical resolution through a surface

configuration of an antecedent-consequent form. A key question emerges from the above:

Is the antecedent-consequent organization a surface relationship, or does it emerge from

the background? If the answer is the former, then to listen to an antecedent-consequent

does not necessarily mean to listen to the background. Alternatively, if the answer is 'it

emerges from the background' then to listen to the antecedent-consequent *is* to listen to the

background.

In this experiment we associate the experience of the interrupted structure to the blockage-

release of blockage image-schema. The establishment of this connection presupposes the

second answer because (i) what is 'blocked' is the structural direction (not the superficial

one) and (ii) if it is not that way, the interrupted structure could be easily elicited by other

image-schematic structures such as the structures that will be used in experiment VII.1.

If we focus on two definitions of the antecedent-consequent function we will find some differences in the conception of the interrupted structure. In the first of them we find that

"... essential to the concept of the period is the idea that a musical unit of partial cadential closure is repeated (italics mine) so as to produce a stronger cadential closure" (Caplin 1998, p. 49).

Two cadences are taken into consideration here: the first cadence is weaker than the second one, and they are linked by the repetition of the musical material. This way of thinking the formal-harmonic structure clearly accounts for a surface phenomenon. According to this definition, instances of interruption in which the antecedent ends in a V-I cadence with I in weak metrical position, can be included as well, assigning a weak character to that cadence.

Concerning the second definition

"...the formal property of antecedence-consequence is, indeed, pertinent wherever partitioning involves a second balancing unit which is in its active elements more decisive in content and in finality of arrival. While other elements often contribute to the expression of antecedence-consequence, the decisive factor in most music is that of tonal orientation, in that [it is] the extraordinary system of relations by which motion is directed toward the first scale degree, toward the tonic triad and toward the primary tonic following conditions of relative fluctuation into secondary, subordinate regions..." (Berry [1966] 1986, p. 402).

Clearly what is important here is the sense of directionality toward the point of greatest stability, and, as a consequence, a definition that lies more on deeper components of the structure, making the idea of blockage - release of blockage more tenable. In this concept of the interrupted structure the features of balance relate more to the resolution of the relationship blockage - removal of restraint than to superficial elements of the musical piece that govern, for example, issues of symmetry or repetition.

VII.3 The force schema and the metaphorical conception of the underlying structure in Schenkerian thought

As was described in detail in Chapter VI, the force image-schema is generated as a consequence of our causal interaction with the environment, and requires the exertion of force as we act upon objects or as we are acted upon by them. According to Johnson (1987) the force schema forms a radial category, that is, a category containing a prototype or most typical member and other less typical members. The prototype is at the centre of the category; the members that are most similar to the prototype are represented as closer to the centre, and those least typical are nearest the outer boundary of the category. Johnson lists the following features of the force schema prototype: 1) interaction or potential interaction between entities; 2) directionality (of a moving object or of the force exerted against it); 3) a single path of motion; 4) origins or sources and targets or goals; 5) degrees of power or intensity; and 6) a structure or sequence of causality. The seven common members of the force schema category are: compulsion, blockage, counterforce, diversion, removal of restraint, enablement, and attraction (see a description of each of them in Chapter VI).

The governing metaphors that Schenker employs in his descriptions of music as a living organism propelled by a life force, arise in the context of his philosophical conception (see Chapters I and VI for a detailed explanation). However, Saslow (1997-1998) succeeds in showing how embodied aspects of knowledge, that is to say, the physical experience of body movement and orientation, might structure Schenkerian concepts themselves and the inferences drawn from them.

Schenkerian statements display many of the features of the force prototype cited by Johnson and exemplify several of the other force schemas of the radial category. Among those features, the more clearly present in *Free Composition* ([1935]-1979) are directionality, a single path of motion, and a source and a goal of motion. Schenker's descriptions of the way force sets into motion the horizontalized version of the chord of nature (*FC*, p 25; p. 4) use biophysical metaphors such as 'comes to life', 'vital natural power', 'motion that grows', 'life impulse', etc. He understands musical motion as a live force that animates the fundamental structure; he contextualizes the idea of motion in terms of travel along a path, a conception that derives from an understanding of time as space. In the fundamental line, motion is the result of the necessity (*Bedürfnis*) of the fundamental passing tones to continue in the same direction.

Schenker strongly posits that there is a goal of this directed motion, stating that "...the fundamental line signifies motion, striving toward a goal, and ultimately the completion of this course" (*FC*, p.4). He also suggests that the direction and achievement of the goal are clearly indicated to the ear, and specifies exactly what the goal of this linear motion must be: the tonic in both voices. "...we feel by nature that the fundamental line must lead downward until it reaches $\hat{1}$ and that the bass must fall back to the fundamental." (*FC*, p.13). In this statement a metaphorical idea of musical understanding is implicitly based, in Lakoff's terms, in our sensitivity to experience the kind of energetics (see Chapter VI for an explanation of the concept of energetics) conveyed by the musical piece, thanks to our capacity to use the emergent features of our live experience and to correspond them to the sound's sequence of the piece in order to *structure* it.

As said above Saslow identified in Schenker's *Free Composition* some of the features Johnson analyses in the force schema prototype. A structure of sequence of causality is

implied, for example, by the descent of the fundamental line. The propulsive force causes descending motion by step from the primary tone to the fundamental tone. The linear progression includes arrival on dissonances that cause the continuation by step in the same direction until the tonic is reached. Different degrees of power might be implied by the relative length of the spans supported by the tones at higher levels (longer spans imply stronger supporting force) or perhaps more convincingly by the strength of the tonic to force other tones to gravitate towards it, as opposed to the weaker force of the dominant. That so many features of Johnson's force prototype, present in *Free Composition*, suggests that Schenker indeed conceived of musical forces along the lines Johnson has posited.

Not only does Schenker's conception of a musical life force exemplify Johnson's force schema prototype, it also presents instances of the compulsion schema, one of the members of the force radial category. In this schema a force interacts with entities that are then compelled into motion in the same direction. Schenker's writings are full of expressions indicating the compulsion of motion toward the ultimate completion of the fundamental structure, such as, for example: the various voices "are forced to move forward" (*FC* p.35). Schenker uses the word necessity (*Bedürfnis*) to express this compulsion. His concept of inner tension also conveys the idea of musical forces. This concept can be explained in terms of several force schemas. First of all, tension (*Spannung*, in Schenkerian vocabulary) is related to the force required by the members of the fundamental line to support, in the architectural sense of pillars, the spans (*Spannungen*) that they control. To the extent that those musical pillars are resisting musical gravity, a counterforce schema seems to be also in operation. The notion of tension also applies when the fundamental line must control the forces that lead the individual tones in many different directions, since each has its own

will. In this sense, the compulsion schema combines with the counterforce schema and the diversion schema.

More important to the present study, Schenker's conception of tension seems to include Johnson's blockage schema. Blockage entails the prevention of a force continuation in a particular direction by an obstacle of some kind. Schenker's notion of tension is considered as deeply rooted in the context of the combat between life and death. Therefore, it is interpreted as being part of the pair *extension* ('movement in the direction of life') and *tension* ('movement in the inverse direction of life') (Kassler, 1983, p.242). The metaphorical concept of tension as movement in the inverse direction of life fits quite well the counterforce and blockage schemas identified by Saslow. According to Schenker, the expansion of the fundamental structure, through transformations from the background to the foreground, implies the entertainment of tension, thus *delaying* the arrival to the goal. This tension (and thus the blockage) is only eliminated when the goal, the final tonic, has been reached and the motion towards it stops. Inherent to the blockage schema is the possibility that encountering an obstacle changes the direction of the force's motion. Since descending motion signifies approaching the ultimate goal (in both parts of the fundamental structure), all upward motion is a diversion or delay of closure. For Schenker, all diversion from motion toward the goal, whether upwards or downwards, is temporary.

> " In the art of music, as in life, motion toward the goal encounters obstacles, reverses, disappointments, and involves great distances, detours, expansions, interpolations, and, in short, retardations of all kinds." (*FC*, p. 5).

This quotation makes the blockage schema explicit. It also brings out the relationship between the compulsion schema and the source-path-goal schema. Both involve motion, a source, a direction, a path, and a goal. Yet, each has other attributes that focus either on the

agent or cause of motion, in the case of the force schema, or the journey itself, in the case of the source-path-goal schema.

To summarize, Schenker's metaphorical understanding of music in terms of a life force conforms to the features of the prototypical force Gestalt as delineated by Johnson. In his conception of musical structure, Schenker employs a number of Johnson's force schemas, the most significant of which for his theory, Saslow believes, are compulsion and blockage. As stated in Chapter VI, image-schemas are flexible and combine each other to form more complex imaginative structures. In the case of the Schenkerian idea of the underlying structure as a life force, it is clear that there is no conflict between the operation of different sub schemas of the force prototype, given that each one highlights a certain aspect of music's unfolding at a particular time.

VII.4 The musical fragments

The 9 musical fragments selected as stimuli are pieces belonging to the repertoire of tonal Western art music (see in Figures VII.2 to VII.10 below the musical fragments used, together with their standard voice-leading analysis). All represent examples of the interrupted form (see a description of the interrupted form above in this chapter, in section VII.2) according to the analyses provided by the literature on music theory (Cadwallader & Gagné, 1998; Salzer, [1962] - 1982; Forte & Gilbert 1982; Schenker [1935] 1979). Each fragment fulfils a whole musical phrase because it is within that span that the interrupted structure takes place. Accordingly, the length of each musical fragment is dependent on the completion of that constraint. There follows a detailed description of each of them.

The first musical fragment (Figure VII.2.a) corresponds to bars 1-16 of the second movement of Beethoven's Piano Sonata, Op.14, No. 1. The duration of the passage is of 15

sec. The passage suggests one of the many possibilities of realization of the fundamental

structure at the foreground level in a musical composition. It presents a

Figure VII.2.a. Beethoven, Piano Sonata, Op.14, No. 1, II, bars 1-16
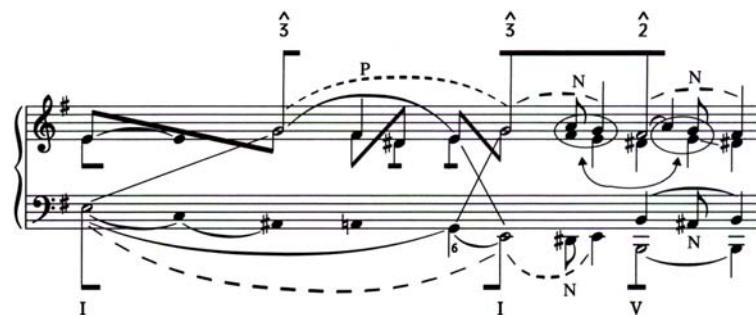
Figure VII.2.b. Beethoven, Piano Sonata, Op.14, No. 1, II, bars 1-8. Voice leading analysis: foreground reduction. (Extracted from Cadwallader and Gagné 1998, p.164)
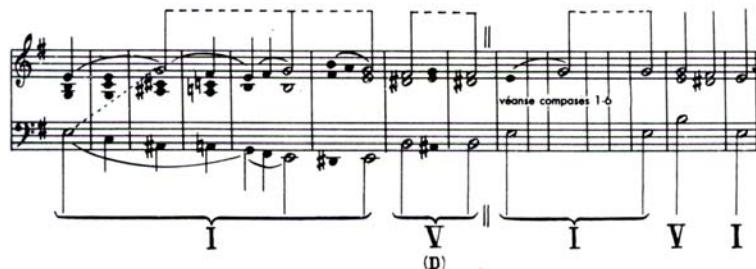
Figure VII.2.c. Beethoven, Piano Sonata, Op.14, No. 1, II, bars 1-16. Reductional analysis of the whole interrupted structure. (Extracted from Salzer [1962] - 1982, vol. 2, ex. 265, p.87)

clear antecedent-consequent phrase with a symmetrical internal division that establishes a

relatively self-contained construction. The articulation of the antecedent occurs in bars 7-8

by means of a half cadence and the full cadence takes place at the end of the consequent in

bars 15-16. The underlying structure of the phrase is a clear example of an interrupted structure of the form $\hat{3}$-$\hat{2}$//$\hat{3}$-$\hat{2}$-$\hat{1}$ (see Figures VII.2.b and VII.2.c). In the first branch of the interrupted structure (Figure VII.2.b) the analysis of the outer voices of the voice-leading shows in the upper voice an initial arpegiation, unfolding an ascendent motion of the third E-G from the inner voice E in bar 1 to G in bar 3, G being the primary note of the *Urlinie*, $\hat{3}$ (the *kopfton*). From G, the linear progression, always in the upper voice, leads downward to F (②) in bar 7, where the form is interrupted. While, in the bass line, the initial tonic is prolonged by means of a transference of register, until V (B) under $\hat{2}$ is reached in bar 7. Due to the initial prolongation of the tonic, the delayed G of the upper voice is structurally associated with E in the bass, being G the upper third of the tonic harmony (e minor), in spite of the diminished seventh chord that appears under G as a foreground harmony in bar 3. The second branch of the interrupted structure begins thematically identical to the first branch, except that the musical material is presented in a different register (in the upper octave), and proceeds similarly until $\hat{1}$ over I is reached by means of a perfect cadence at bar16.

Figure VII.3.a. Beethoven, Piano Sonata, Op.2, No. 1, II, bars 1-8

Figure VII.3.b. Beethoven, Piano Sonata, Op.2, No. 1, II. Voice leading analysis: foreground reduction. (Extracted from Cadwallader and Gagné 1998, p.171 and p.173)

The second musical fragment (Figure VII.3.a) corresponds to bars 1-8 of the second movement of Beethoven's Piano Sonata, Op.2, No. 1. The duration of the passage is 59 sec. Overall, the opening of this Adagio is similar to the opening of Beethoven's Op. 14, 1, II previously analysed, in thay way the interrupted structure is unfolded. In this case, an eight-bar structure of a symmetric antecedent-consequent is displayed along a longer time span. The interruption also obeys the form $\hat{3}$-$\hat{2}$//$\hat{3}$-$\hat{2}$-$\hat{1}$. However, differences can be identified in the details of the musical surface, accounting for the fact that two very different musical surfaces can be based on identical deep middleground structures. The antecedent presents an interesting case of the realization of a voice-leading procedure in the upper voice that unfolds the primary tone A in bar 1, by means of a linear technique of superposition that involves the transfer of inner-voice tones to a higher position in the register, resulting in A-C-B bemol-A (see Figure VII.3.c); the enlarged repetition of this

285

melodic unfolding takes place afterwards at a larger scale in bar 3, before reaching G (2) at

bar 4 in the end of the first branch of the interrupted structure. Thus, the primary tone A is

prolonged from bar 1 to bar 3. The prolongational procedure is completed with a

movement into an inner voice from A to F in bar 3. This third-motion A F occurs in the

Figure VII.3.c. Beethoven, Piano Sonata, Op.2, No. 1, II. Foreground reduction showing the reaching over procedure. (Extracted from Cadwallader and Gagné 1998, p.171 and p.173)

context of a prolonged tonic harmony in the bass, and therefore associates the prolonged $\hat{3}$

in the upper voice with the prolongation of the tonic chord. Interesting to the purposes of

this thesis is the comment of Cadwallader about this passage:

> "This is a good example of why Schenker often used organic metaphors to describe the
>
> processes of tonal music: the shape of the upper voice seems to 'grow' from bar 1, which we
>
> might metaphorically regard as a musical 'seed'."(Cadwallader and Gagné 1998, p. 170).

Concerning the consequent phrase, we find a surface difference at the beginning: now the initial ascending leap of a sixth from the anacrusis C to the downbeat A of bar 5 is filled in with diatonic and chromatic steps; although C belongs to the second branch, it functions somehow as a link between the two phrases of the interrupted form. The second difference is found in the expanded use of the register in bar 6, which is realized by means of the fourth progression C-D-E-F that had been previously presented in the first phrase at a lower register. This arrangement corresponds to what Schenker calls 'motivic parallelism' (see motivic parallelism in Chapters II and VI). The use of the 'reaching over' procedure in bar 7 resembles the melodic elaboration of the upper voice in bars 1 and 5. However, in spite of the apparent surface differences between both branches, they are strinkingly similar at deeper levels of the underlying structure of this passage.

The third musical fragment presents the beginning of the second movement of Beethoven's Piano Sonata Op 10, 1. The duration of the passage is 52 sec. It is a sixteen-bar pasage that shows another example of the fundamental interrupted structure. Each one of the two eigt-bar phrases of the whole structure is divided into two four-bar half-phrases. Given the length of the whole passage and due to its very slow tempo, the surface harmonies found at bars 4, 8 and 12 contribute to generate phrasal articulations in those points of the musical discourse. In this fragment, it is apparent how the form is composed-out by means of the arrangement that emerges from the harmonic and melodic prolongations. The antecedent phrase presents an initial ascent, that is to say, a stepwise line that ascends from A to C reaching the primary tone ($\hat{3}$) at bar 5 (see Figure VII.4.b).

Figure VII.4.a. Beethoven, Piano Sonata, Op.10, No. 1, II, bars 1-16.

Figure VII.4.b. Beethoven, Piano Sonata, Op.10, No. 1, II. Voice leading analysis: foreground reduction. (Extracted from Cadwallader and Gagné 1998, p.178 and p.182).
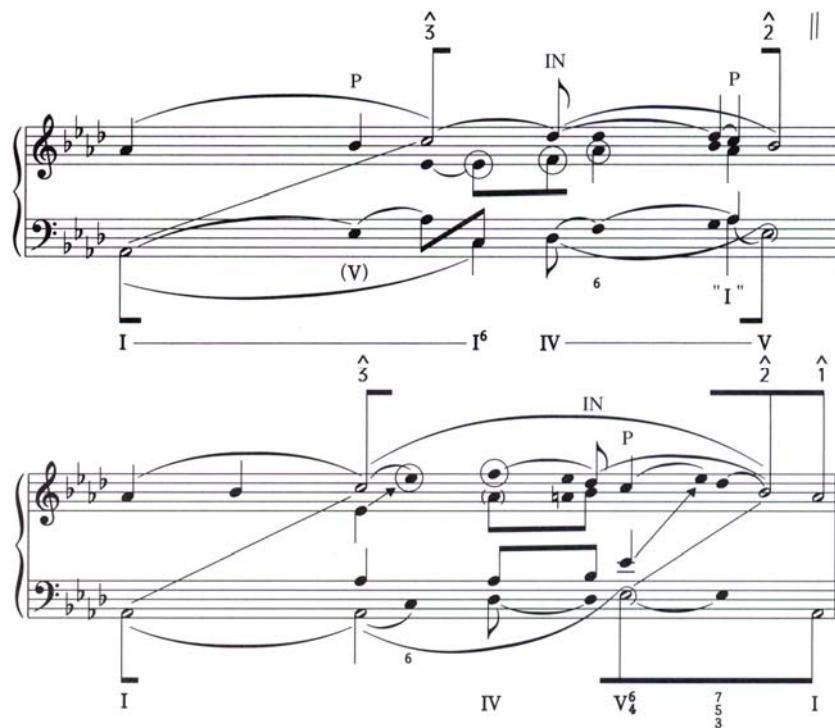
Figure VII.4.c. Beethoven, Piano Sonata, Op.10, No. 1, II. Voice leading analysis: middleground reduction. (Extracted from Cadwallader and Gagné 1998, p. 178 and p. 182).

The role of the V-I resolution in bars 4-5 is to contribute to the development of a broader harmonic motion that provides cohesion to the antecedent phrase. By prolonging I beyond the end of bar 4 the composing-out of the form builds a bridge between the first subphrase and the second one, unifying the whole antecedent. Once C (3) has been reached at the beginning of bar 5, it moves to D in bar 6, and comes back to C in the second beat of bar 7, before descending to B bemol in bar 8.

It is interesting to point out here the expansional process of the second subphrase of the antecedent, where the inner voice in the alto register is duplicated in octaves above C by means of a superposition over the primary tone that creates expansion and content in the unfolding of the musical phrase. Thus, the harmonic path towards IV and, finally, the arrival at V at the end of the antecedent phrase, is composed-out by means of an expanded cadential six-four that is located at the first beat of bar 8. The movement of the bass

expands the subdominant harmony by means of a motion from the root to the fifth of the subdominant chord, which in turn is the root of the tonic chord. This movement, that creates a link between IV and I, serves as an expansion of the subdominant harmony that prolongs it until V is reached, at the end of the antecedent.

After the interruption at bar 8, the consequent begins at bar 9, presenting the same musical material as in bars 1-4, but with different figuration in the bass. The final subphrase of the consequent achieves completion recomposing the musical material by means of three main procedures: by composing out the inner-voice notes that go over C as grace notes that precede the notes at the higher register, by expanding the subdominant harmony by means of a IV II6 motion in a shorter temporal span, and by recalling the initial motivic elaboration of the upper voice at the beginning of the antecedent in the final cadence.

The fourth musical fragment is Chopin's Prelude Op 28, 1 (Figure VII.5.a). The duration of the piece is 30 sec. This musical piece constitutes an example of an interrupted structure that elaborates a chorale texture (Forte & Gilbert 1982). The analysis of the upper voice shows the presence of a recurrent 5-6 motivic pattern. The prelude represents also an interrupted form of the type $\hat{3}$-$\hat{2}$//$\hat{3}$-$\hat{2}$-$\hat{1}$, but the salience of $\hat{3}$ as the primary tone is not as clear at the beginning as in the previous Beethoven examples.

The first appearance of E ($\hat{3}$) occurs in bar 5 as an appoggiatura over a subdominant surface harmony that operates as an extension of the tonic harmony, before progressing to 2 over V in bars 6-8, where the interruption takes place (see Figure VII.5.b). The second term of the interruption begins in bar 9, unfolding the form by means of a registral enhancement from bar 17 to bar 21, in which the upper voice reaches C in the upper octave, and then turns downwards until D is reached in bar 24, to finally arrive at C over I 6/4-5/3 in bar 29, completing the form by a sustained pedal until the end of the piece.

Figure VII.5.a. Chopin, Prelude Op.28, No. 1.

Figure VII.5.b. Chopin, Prelude Op.28, No. 1.Reductional analysis of the complete interrumpted structure (Adapted from Forte & Gilbert[1962] 1982, p. 193 and p. 194).

Figure VII.6.a. Chopin, Prelude Op.28, No. 3.

Figure VII.6.a. Chopin, Prelude Op.28, No. 3 (continuation).

The next musical fragment is Chopin's Prelude Op 28, 3 (Figure VII.6.a). Its duration is 48 sec. In this piece, that represents another example of the $\hat{3}$-$\hat{2}$//$\hat{3}$-$\hat{2}$-$\hat{1}$ interrupted form, the primary tone B appears clearly at bar 5, after an ascending gesture that registrally expands the consonant leap G B over 4 octaves (see Figure VII.6.b). From bar 6 onwards the upper voice progresses to $\hat{2}$, appearing first at bar 8 as a resolution of the secondary dominant of bar 7. This situation is repeated in bars 9 and 10. Chopin's surface elaboration prolongs $\hat{2}$

293

from its appearance until it reaches V in bar 10. The first branch of the interruption is completed in bar 11, presenting C at the upper voice as a surface seventh of the V harmony in the upper register. This tone will reappear in the second branch as an incomplete neighbour of B in bar 20, prolonging the sonority of the subdominant harmony and descending an octave in the upper voice to gain $\hat{3}$ at bar 22, leading to $\hat{2}$ at bar 23 and finally arriving to $\hat{1}$ at bars 24-26.

Figure VII.6.b. Chopin, Prelude Op.28, No. 3.Middleground reduction. Antecedent phrase. (Extracted from Forte & Gilbert [1962] 1982, p.203).
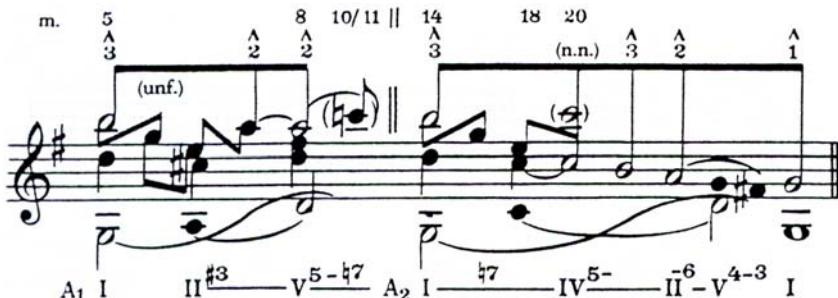
Figure VII.6.c. Chopin, Prelude Op.28, No. 3. Reductional analysis of the complete interrupted form. (Extracted from Schenker [1935] 1979, fig. 76-2).

Musical fragment 6 is the second movement of Mozart's Piano Sonata K. 311, whose duration is 20 sec (Figure VII.7.a). The first branch of this interrupted structure presents the primary tone B in bar 1, at the beginning of the passage, and prolongs I during the four initial bars, by means of a neighbour surface V6/5 chord, that has C in the upper voice (see Figure VII.7.b). After returning to B, by means of an exchange of voices in bar 3, the

melody climbs upwards reaching E, and descends afterwards until $\hat{2}$ over V is reached in bar 4, where the form is interrupted. The second branch is almost similar to the first one, except for slightly different figurations in the upper melody and the appearance of the final cadence one beat earlier than before.

Figure VII.7.a. Mozart Piano Sonata K 311, II
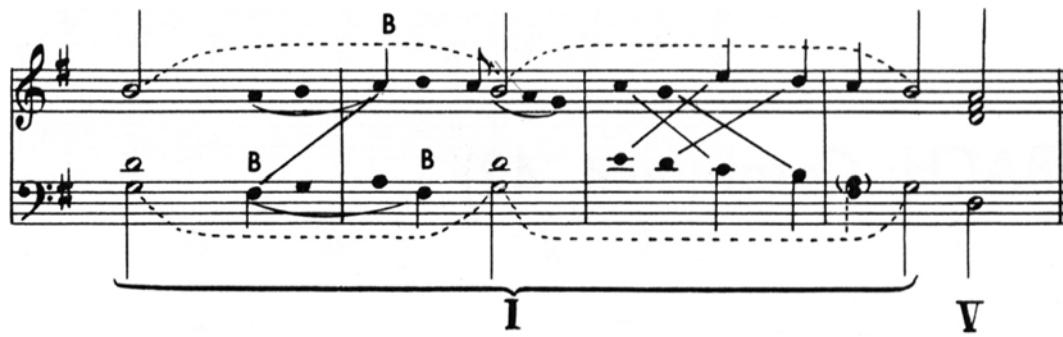
Figure VII.7.b. Mozart Piano Sonata K 311, II. Foreground reduction of the first branch (Extracted from Salzer[1962] - 1982, p. 50, ex. 183).

The seventh musical fragment is the first section of Mozart's Piano Sonata, K 545, second movement (Figure VII.8.a).The duration of this fragment is 1 minute. The sixteen first bars

of this passage present a voice-leading unfolding of the interrupted form $\hat{5}$-$\hat{4}$-$\hat{3}$-$\hat{2}$//$\hat{5}$-$\hat{4}$-$\hat{3}$-$\hat{2}$-$\hat{1}$.

The whole passage shows an apparently simple design. Once the voice-leading is scanned

in more detail, it reveals an elaborated motivic arrangement. The

Figure VII.8.a. Mozart Piano Sonata K 545, II

Figure VII.8.b. Mozart Piano Sonata K 545, II. Foreground reduction. (Extracted from Beach 1984, p. 229).

opening two bars of the antecedent unfold an ascending line in the upper voice, that goes from B to the primary tone D ($\hat{3}$) in bar 2 (see Figure VII.8.b). A turn figure decorates the passing tone C between B and D. After D is reached, a falling arpegiated fifth takes place. The entire melodic phrase is elaborated on the basis of this falling fifth. The primary tone D ($\hat{5}$) is the first tone of the falling arpegiated fifth. From then onwards, several varied presentations of this melodic figure lead the development of the upper line, prolonging $\hat{5}$ until $\hat{4}$ over IV is reached in bar 7; finally, $\hat{2}$ over V is arrived at in bar 8, where the form is interrupted. The consequent phrase presents a varied repetition of the musical content of the antecedent, mainly by means of an increment of the ornamental density of the melodic surface. The final cadence occurs at bars 15-16. Its voice-leading arrangement reveals a common compositional procedure of Mozart (Beach, 1984) that consists in delaying the occurrence of $\hat{3}$, that instead of sounding over the cadential 6/4, appears in the following beat as part of the turn around $\hat{2}$ over V.

Figure VII.9.a. Schubert Impromptu in Gb major op 90, No. 3. Bars 1-8.

Musical fragment eight is Schubert's Impromptu in G bemol major op 90, 3 (Figure VII.9.a). Duration: 32 sec. The fragment presents an example of voice leading motion of the upper voice into an inner voice. The interrupted form shows a clear $\hat{3}$-$\hat{2}$//$\hat{3}$-$\hat{2}$-$\hat{1}$ structure. The primary tone B bemol ($\hat{3}$) is prolonged along bars 1-3 and the first part of bar 4, by means of two motions into an inner voice: (i) a consonant leap from B bemol to G bemol in bar 2 and (ii) a more local level prolongation of A bemol, through a fifth descending motion into an inner voice to D bemol

Figure VII.9.b. Schubert Impromptu in Gb major op 90, No. 3.Foreground and middleground reductions of the antecedent phrase. (Extracted from Cadwallader and Gagné 1998, p.133).

(see Figure VII.9.b) After the prolongation of A bemol, the main upper voice reaches G bemol in bar 4, and goes up again to the *kopfton* 3 (B bemol) in the second beat of bar 4,

before progressing to $\hat{2}$ (A bemol) in the final part of bar 4, where the interruption occurs. The second branch of the interrupted structure unfolds in a similar way, by means of two motions into an inner voice, the first in half-notes between B bemol and E bemol, and then up to A bemol, which is prolonged over II and V in quarter-notes, until it resolves in bar 8.

The last musical fragment is Schubert's Impromptu in B bemol major, Op. 142, No. 3 (Figure VII.10.a). The fragment has a duration of 24 sec. In the antecedent phrase, the upper melody unfolds a two-voice characteristic pattern, outlining a top - inner - top movement (Cadwallader and Gagné, 1998) that prolongs the primary tone $\hat{3}$ (D) throughout bars 1-2. The bass also shows a similar treatment that complements that of the upper line (see Figure VII.10.b).

Figure VII.10.a. Schubert Impromptu in Bb major, Op. 142, No. 3.Bars 1-8.

The second part of the antecedent phrase shows some displacements of tones in the foreground. In bars 3-4, an unfolding procedure displaces C in the upper voice as the resolution of D over II6/5, to finally imply it over V in the second part of bar 4, where the interruption takes place. The consequent phrase presents a very similar unfolding, with the only surface exceptions of the change of register of the whole phrase, and a change of direction in the upper voice over II, before the occurrence of the final cadence.

Figure VII.10.b Schubert's Impromptu in Bb major, Op. 142, No. 3. Foreground reduction of the antecedent phrase. (Extracted from Cadwallader and Gagné 1998, p. 130).

VII.5 The musical reductions

Three musical reductions were used as stimuli: voice-leading, contour and rhythm reductions. Concerning the voice-leading reductions, the problems found with the analyses provided by the differences sources of musicological literature were that they happened to be the result of following rather different compositional criteria to perform the analyses; that is to say, some of them showed a more contrapuntal conception and others were more harmonic and homophonic, independently of the textural nature of the original musical fragment. Besides, it was not always possible to find the reductional analysis of the complete interrupted form (sometimes only the reduction of the antecedent phrase was

provided) and finally in some cases the reductional analyses corresponded to different reductional levels. Therefore, it was decided to compose the voice-leading reductions of the nine musical fragments. The analyses provided by the different sources were adapted, trying to produce, where possible, an outlook as similar as possible between the different reductions. A foreground reductional level was used in the elaboration of the reductions in order to keep, to a certain extent, the rhythmic pace of the musical surface.

Concerning the melodic contour reductions, they were fully created. They presented the contour reduction of the upper melodic voice. Some of the rhythmic attributes were kept, in order to provide a minimal sense of ongoing movement, and others were reduced. The notes included in the reductions kept the rhythmic values of the notes in the original fragment, except for some situations in which a normalizing procedure - that displaced the notes toward stronger metrical positions - was run, in order to provide the fragment with a sense of musical coherence. In the melodic contour reductions the register was kept as in the original, that is to say, it was not reduced.

Finally, the rhythmic reductions presented the rhythm of the melodic line using only one tone: the tonic note.

The length of all the musical reductions was equal to the length of the corresponding musical fragment. Only the rhythmic reductions kept the whole rhythm of the melody and 'lost' other attributes such as the melodic contour.

Figures VII.11 to VII.19 show the three musical reductions of the nine musical fragments that were used in the experiment.

Figure VII.11. Contour, Voice-leading and rhythm reductions of Beethoven Op. 2, 1, II, used in the experiment.

Figure VII.12. Contour, voice-leading and rhythm reductions of Mozart K.311, II, used in the experiment.

Figure VII.13. Contour, Voice-leading and rhythm reductions of Beethoven Op. 14, No.1, II, used in the experiment.

Figure VII.14. Contour, voice-leading and rhythm reductions of and Schubert,Op. 90, No. 3, used in the experiment.

Figure VII.15. Contour, voice-leading and rhythm reductions of Mozart K, 545, II, used in the experiment.

Figure VII.16. Contour, voice-leading and rhythm reductions of Beethoven Op. 10, No.1, II, used in the experiment.

Figure VII.17. Contour, voice-leading and rhythm reductions of Chopin, Op. 28, 1 used in the experiment.

Figure VII.18. Contour, voice-leading and rhythm reductions of Chopin, Op. 28, 3 used in the experiment.

Figure VII.19. Contour, voice-leading and rhythm reductions of Schubert, Op. 142, No.3 used in the experiment.

VII.6 The visual animations

The visual animations that were used in the experiment are relatively neutral representations of the main features of the proposed image-schemas[1]. They operate as visual primings (P) of image-schematic representations. P1 is a visual animation that is assumed to activate the blockage-removal of restraint image-schema. In it, a ball moves

---

[1] The visual animations were elaborated by Andres Riva using Adobe After Effects.

exerting force in a determined direction and then faces an obstacle that blocks the ball's motion; afterwards, the ball tries to overcome the blockage, impulsing itself to reinitiate the movement, in order to finally succeed in overcoming that blockage (Figure VII 20 shows a fixed representation of the visual animation). P2 and P3 are two different visual animations. P2 consists of a ball moving alternatively up and down, and P3 is a ball flashing at different time spans (see in Figures VII.21 and VII.22 fixed representations of the alternative visual animations). They are assumed not to activate the blockage - removal of restraint image-schema.

Here it is posited that these kinetic images will activate different image-schemas, which are themselves amodal, because they are abstractions of visual, aural, kynaesthetic and/or physical relationships (see Johnson, 1987). Once activated, the image-schemas will allow, by means of a cross-domain mapping process, the attribution of meaning to rich information in another domain, in this case, the musical domain.
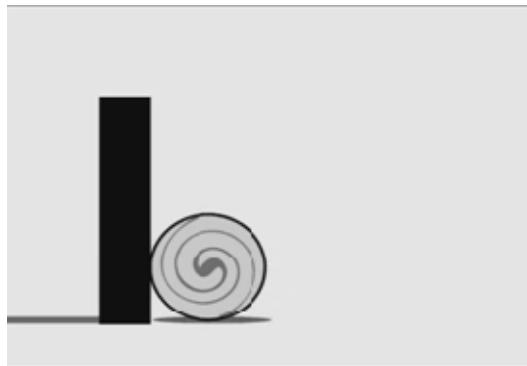
Figure VII.20. Fixed representation of the visual animation corresponding to the activation of blockage-removal of restraint image-schema.
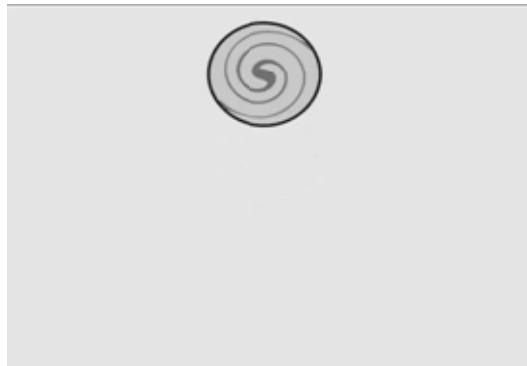
Figure VII.21. Fixed representation of the visual animation corresponding to a ball moving alternatively up and down.

Figure VII.22. Fixed representation of the visual animation corresponding to a ball flashing alternatively at different time spans.

VII.7 Rationale

In the current study, the listener's experience of the Schenkerian interrupted structure is tested. It is assumed that the listener creates meaning while experiencing, consciously or unconsciously, the interrupted structure in terms of the structural metaphor blockage-removal of restraint (Saslaw, 1997-1998).

In the experiment, each melodic fragment is contrasted with different musical reductions of that fragment. In the context of this experiment, the concept of reduction is understood in a

wider sense, not just referring to the standard graphic devices of Schenkerian theory or their derivations of *GTTM* or *IR* theories, but as a *simplification* of the surface level of the piece. In each reduction some structural properties are kept while others are dismissed. A priming paradigm is used in the experiment. The visual animations activate in the participants different image-schemas, on the assumption that they will be used to perform the comparisons between the musical fragments and the musical reductions by means of a process of cross-domain mapping. In the cross-domain mapping process they will hear A in terms of B, being A the musical fragment of the tonal piece and B the structural features highlighted in the musical reductions. It is assumed that, to the extent that the primed image-schema correlates with the characteristic structural feature that is highlighted in the musical reduction, the matching between musical fragment and musical reduction is going to be higher than in the case in which the activated image-schema does not correspond with the structural feature highlighted by the musical reduction.

Given that each musical reduction highlights a particular structural feature of the musical fragment, we assume that the musical piece is heard in terms of all the structural features highlighted by the musical reductions. The key issue is that each visual priming is assumed to activate a determined image-schema that elicits in the listener a tendency *to hear more in terms of one structural feature than of the others*. Therefore, through a process of cross-domain mapping, the participants will tend to estimate higher the correspondence between the musical fragment and the reduction that conveys the structural feature that is assumed correspond to the image-schema activated by the visual priming.

VII.7.1 Hypotheses

General hypothesis: The prolongational structure is understood in terms of some structural metaphor.

The specific hypothesis of the current experiment can be formulated as follows: the prolongational structure of the type interrupted form has cognitive reality at the level of the musical phrase, to the extent that it is metaphorically understood in terms of the subtype of the force image-schema blockage-removal of restraint, as explained by Johnson (1987).

VII.7.2 Prediction

When an image-schema of force is activated with a blockage-release of blockage visual priming, the musical reduction of the interrupted structure will match more closely with the musical fragment than any of the other musical reductions. When this force image-schema is not activated (because a different visual priming is used instead) the voice leading reduction will be less closely matched with the musical fragment.

VII.8 Method

VII.8.1 Subjects

Thirty-one professional musicians, average age 29 years, average musical experience 16 years, volunteered to participate in the experiment. Criteria for selection of this population were similar to the criteria exposed in Experiments V.1 to V.4 (see Chapter V).

VII.8.2 Stimuli

(i) 3 Visual animations that are relatively abstract representations of the main features of the proposed image-schemas were used as visual primings (P). P1 is a visual animation

that is assumed to activate the blockage-removal of restraint image-schema. P2 are two different visual animations that are assumed to do not activate the blockage-release of blockage image-schema (see section VII.6 for a detailed explanation of the visual primings). Each of the animations was of 10 seconds duration.

(ii) 9 fragments of musical pieces (M) belonging to the repertoire of tonal Western art music were used (see section VII.4 for a detailed explanation of the musical fragments used as stimuli).

iii) Musical reductions (R) that highlight different structural features of each musical fragment were used. R1 are voice-leading reductions; R2 are non voice-leading reductions (see section VII.5 for a detailed explanation of the composition of the musical reductions used in the experiment).

VII.8.3 Apparatus

The experiment was designed and run using the experimental software DirectRT from Empirisoft Co.NYC.

VII.8.4 Procedure

In the first part of the experiment, the participants completed a tutorial session containing instructions about the organization of the experimental trial, where they were able to rehearse the experimental task. Information about what a musical reduction consisted of was also provided. The visual animations were described as fragments of visual information that were used to separate the musical trials and to help participants to concentrate in the music that followed, but to which they had to attend, nevertheless. The final instruction about the resolution of the task required the subjects to follow this order: i) to look at the visual animation, ii) to listen to the musical fragment, iii) to listen to the

musical reduction, iv) as soon as the sound signal is heard to press a *Yes* key to say that the reduction *does match*, or a *No* key to say that the reduction *does not match* the musical fragment and v) to press another key to say how sure he/she was of the answer, using a 9 points scale that ranged from 1 not sure to 9 sure.

Subjects were tested in two individual experimental sessions.

VII.8.5 Experimental Design

Each trial consisted of the following sequence:

P (visual priming) - M (Musical fragment) - R (musical reduction) - Sound signal - Time to answer.

The test contained 63 trials, formed by combinations of the 9 musical fragments, according to their agreement/disagreement between priming and reduction. The relationship between visual priming and musical reduction resulted in the following combinations of the different experimental trials:

(i) Match between priming and reduction: P1 (blockage-removal of restraint) - R1 (voice-leading reduction) - 9 examples -; P2 (non blockage-removal of restraint) - R2 (non voice leading reduction) - 9 examples -.

(ii) No match between priming and reduction: P1–R2 (18 examples); P2–R1 (18 examples).

9 extra sequences with foil reductions were included to elicit continuous attention to the musical unfolding of each reduction, in order to produce a valid goodness-of-fit response. Trials were randomised in such a way that each participant listened to a different order of presentation of the examples.

VII.8.6 Results

The goodness-of-fit responses were converted into an 18 degree scale that ranged from 1 do not match / sure to 9 do not match / unsure - 10 match / unsure to 18 match / sure. The means of the responses to the different combinations of visual priming - musical reduction were obtained.

In order to see if participants metaphorically understood the interrupted structure in terms of the blockage-removal of restraint image-schema, two groups of responses could be compared:

Voice Leading Reduction (R1) – Blockage-Removal of restraint visual priming (P1)

Voice Leading Reduction (R1) – Other visual primings (P2: up-down; flashing beats).

However, if differences were found between those two groups of responses, it might be possible that they were not an outcome of the association between reduction and visual priming, but just the result of a pure priming effect. In other words, the differences could reflect the priming effect *per se*. Therefore, it was necessary to have a group of musical examples, in order to contrast those two different primings with other reductions. If no differences were found between the two types of priming in those musical examples, then it would be possible to assert that the differences found in the previous groups were due to the association between priming and reduction and not to the effect of the priming per se. Consequently, the means of the goodness-of-fit responses to the four combinations of priming-reduction were compared: P1-R1; P2-R2 (match between priming and musical reduction); P1-R2; P2-R1 (no match between priming and musical reduction).
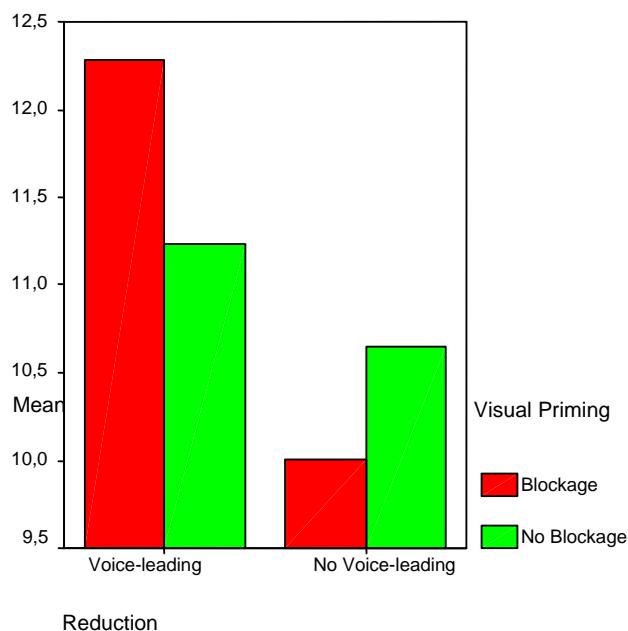
Figure VII.19. Means subject'responses to the four combinations of Priming (Blockage-release of blockage/Different priming) and Reduction (Voice leading reduction / Other reductions).

The results are shown in Figure VII.19.

A repeated-measures ANOVA, with 9 Musical Fragments x 2 Musical Reductions (Voice leading reduction / No voice leading reductions) x 2 Visual Primings (Blockage-removal of restraint/No blockage-removal of restraint) as factors was run.

Factor Visual Priming was not significant, meaning that results are not due to its unique effect. Factor Musical Reduction was significant ($F [1, 30] = 14.947$; $p < .001$). It can be observed that the matching between the voice leading reductions and the musical fragments is estimated higher than the matching between the fragments and the non voice-leading reductions. This result could be informative of potential differences in the compositionality of the different reductions. It seems as if, overall, the voice-leading

reduction conveys a more consistent representation as a simplification of the musical fragment than the other two non voice leading reductions.

Most important to our purposes is that the interaction between Reduction and Visual Priming was significant ($F [1, 30] = 7.608$; $p < .01$). The results confirm the prediction: when subjects are primed with the blockage-removal of restraint visual animation, they estimate the association between the musical fragment and its voice leading reduction higher; and conversely, when they are primed with different visual animations, they estimate the association between those fragments and their voice leading reductions lower.

VII.9 Discussion

These results support the general hypothesis that metaphorical thinking shapes music experience, and the specific hypothesis that the listener understands the interrupted structure of tonal compositions in metaphorical terms. Once an image-schema has been activated, a cross-domain mapping process takes place and the participant uses knowledge from one domain to understand knowledge of another domain; in this case, they use the image-schematic knowledge of force relationships in order to hear A in terms of B, A being the musical fragment and B the underlying voice-leading of the interrupted structure, highlighted by the musical reduction. When the image-schema that has been primed maps with the structural feature that is conveyed by the reduction, the match between reduction and musical fragment is higher as compared with the case in which the activated image schema does not map with the attribute highlighted in the musical reduction.

According to our results, structural metaphors seem to be more than just linguistic constructions used in the theoretical discourse about music. They appear to operate as internalized models of cognitive processing that listeners activate during the experience of

attending to aspects of the musical structure. That is to say, structural metaphors are at the core of the idealized cognitive models that participants activate and use to perform categorizations of sounds in terms of structural features.

The approach developed by Lakoff & Johnson (1999) relative to the nature of metaphorical knowledge as a factor that shapes our experience supports hypotheses of the experience of musical structure in metaphorical terms. Structural metaphors are based on systematic correlations between different domains of human experience. They influence the way meaning is assigned to action. Their properties form groups that operate as wholes, or put it differently, as unitary structures that can be understood as one-to-many correspondences. They are applied automatically and regularly to assess aspects of reality. They enhance the prototypical nature of human basic knowledge, and connect, by means of mapping processes, operations between different cognitive domains.

Structural coherence, a feature that is highlighted in the analysis of tonal music, can be understood in the context of the cross-domain mapping procedure as a metaphorical experience that occurs when a person is capable of overlapping a multidimensional structure of elements and/or properties of an object onto the structure that corresponds to another object. Taking the prolongational structure in music as an example of analysis, tonal coherence occurs if the listener is capable of categorizing sounds in terms of prolongational features (Chapter VI details some of the main analytical procedures employed by Schenkerian theory to approach the prolongational structure of tonal music).

In the case of the current experiment, structural coherence, experienced as a metaphorical process, occurs when the listener is capable of understanding the harmonic-melodic process of tension that unfolds the relationship $\hat{3}$-$\hat{2}$//$\hat{3}$-$\hat{2}$-$\hat{1}$ in the upper voice, and I-V//I-V-I in the bass, as purposeful movement that is blocked and subsequently released, to finally

arrive at the resolution of the tension that was interrupted in the middle of that formal phrase division.

If theories relative to the principles of musical structure provide assumptions about certain phenomena of music cognition, then it is the work of the psychology of music to derive formulations that are demonstrable by means of experimentation. The results reported here support the hypothesis of the prolongational structure as an *idealized cognitive model* (Lakoff, 1987) that is substantiated through a cross-domain mapping process of categorization.

In this thesis it was considered that the underlying structure, developed in Schenkerian theory as an imaginative metaphor, has a descriptive value, whose cognitive status needed to be explored. The present research tried to fulfil this purpose. If music is essentially a kind of activity that, in words of Cross (2005) has the property of "*embodying*, *entraining*, and *transponsably intentionalizing* time in sound and action" (p.35), it was considered fruitful to study the relationship between those aspects of the embodied experience as a factor that activates imagination. It is also remarkable that in the core of Schenker's seminal idea of organicism, developed in his metaphorical descriptions of the underlying musical structure, is hidden the experientialist approach of embodied cognition, including the force metaphor tested in this experiment.

The voice leading reductions used in experiment VII.1 were composed following principles that acknowledge a long tradition in the practice of music analysis. According to the results, those reductions provided consistent examples of the attribute that was being studied. The contrasting reductions used in the experimental design included cases of different reductions that were treated overall as examples of non-voice leading reductions. The methodology employed in the experiment, which included the priming of image

schematic representations and the matching between musical reductions and musical fragments, using the activated image schema, proved to be a valid experimental tool to study the imaginative nature of music cognition. It could be an interesting continuation of this inquiry to explore the cognitive status of other image schematic representations. In order to do this, it is necessary to explore more thoroughly the composition of alternative musical reductions, in order to produce examples as consistent in the communication of the highlighted attribute as the examples provided by the voice leading reductions.

CHAPTER VIII

GENERAL DISCUSSION

> *"A better account of the relationship between perception and music theory is needed if the discipline is not to collapse into the history and sociology of aesthetics on the one hand, and the psychology and pedagogy of note-to-note structure on the other"*
> (Cook, 1989, p. 123).

VIII. 1 Introduction

This thesis investigated the status of prolongational structure as experienced in cognition.

The concept of prolongation, as described in the theories of underlying structure, refers to a structural phenomenon that involves a group of elaborative processes in which some pitch events, such as chords and notes, which are structurally more stable, are assumed to remain active within the musical stream of subsidiary events that elaborate them, even though they have no physical substance. The pervasive use of this concept in different musicological sources suggests its value as an analytical / theoretical tool to be applied to the study of the underlying structure in tonal music and, at the same time, elicits interest in its psychological implications.

Prolongation was described for the first time at the beginning of the twentieth century in Schenkerian theory. Since then, it has been part of the music-theoretical developments throughout the past century, and was also used in new theories of musical structure that emerged from the influence of cognitive science and linguistics during the second half of the twentieth century, such as *GTTM* by Lerdahl and Jackendoff and *I-R* by Narmour.

Overall, prolongation conveys the idea of a one-to-many relationship between pitch events in tonal pieces. But, nevertheless, the ways such a one-to-many relationship is interpreted in the diverse theoretical sources seem to obey rather different, if sometimes conflicting ontologies.

The representation of the underlying structure can be understood according to alternative modes of abstraction. Some reductional representations are seen as quite feature-specific, in the sense that they keep similarities with the audible parameters of the musical piece, while others are not as audible in terms of their attribute-specificity. In the first case, abstractions are like copies of specific events of the musical piece with their corresponding features of frequency and location in the musical score; in the second case, on the other hand, they are abstractions of a different kind; for example they are rhetorical and/or schematic; they are not like literal piece events. Referring to this issue, Benjamin (1982) suggests: "one cannot hear a transition or a climax; rather, one hears a succession of audible events *as a* transition or some event *as* climactic" (my italics, p. 29). But also: "…the fact remains, however, that functional labels allude *metaphorically* to relations between things with specific audible properties. What exactly is a "prolonged dominant" if it is not something with specific pitch content?" (my italics, p. 30).

In spite of the indisputable sonic parametric substance that underpins a pitch event, these two different ways of conceiving the underlying representation of a tonal piece are manifestations of two rather different conceptions of what an abstraction of an underlying event consists of. In one conception, the basic hierarchical relationship that takes place between the events of a given passage is that in which one of them, which is experienced as the most focal or stable, is chosen to represent other events that are assumed to be related to the former; thus, a *copy* of this event is borrowed from the piece and appears at a

higher level of representation, in which such a copy is considered to represent the passage as a whole. This procedure is repeated with the subsequent passages all along the concatenated organization of the musical piece. In the other conception, the structural event is considered to be somehow manifest throughout the passage, taking the form of some structural feature, for example an arpeggiation, a linear progression, an interruption, etc. In this case, as stated earlier, the relationship between the piece and the underlying representation takes the form of a symbolic abstraction that represents the underlying unity and simplicity of the multiplicity of events of that musical passage.

Following from this, the treatment of prolongation in the theoretical sources reviewed in this thesis revealed different interpretations of a "nominally identical" concept. It was proposed here that these rather alternative, if not conflicting, interpretations of the abstraction of the prolongational structure might be the result of the tension that emerges from different epistemes to which the analysis of the underlying hierarchy in music could be ascribed. It was an assumption of this thesis that these two rather differing ontologies reflect, in part, the route followed by the concept of prolongation throughout the twentieth century. This analysis includes the history of the intersections between the fields of music theory and psychology of music, from which some of the most remarkable and controversial approaches to this topic emerged (Cook, 1989; 1990; Narmour, 1977; Lerdahl & Jackendoff, 1983; Benjamin, 1982; Cohn and Dempster, 1992; Larson, 1997; Keiler, 1978, among others).

The concept of prolongation, as stated above, originates in the context of Schenkerian theory; its musicological origin sets a close relationship between the concept and the practice of music analysis; and more recently, prolongation emerges reframed in those musical theories that apply the assumptions of cognitive science to musical modelling,

with its strict rules of recursivity and disjointedness. The above-mentioned epistemological tension is evident in the efforts of some of the most outstanding representatives of this tendency to locate prolongation within a conceptual framework that may not be easy to sustain. The problem with this later course of action is that some important things about prolongation were omitted.

First, it is apparent from Schenkerian practice that prolongation is not a concept that emerges spontaneously in the musical metalanguage of the everyday, amateur listener; its cognitive status as a representation is instead linked to the praxis of the analyst (Kielian-Gilbert; 2003; Cook, 1989). By means of the application of analytical procedures, a Schenkerian *reading* is obtained, and is understood as the product of the speculative activity in which the musical text (the musical *score*) is 'scanned' in order to derive the analytical solution, in a way that is similar to the ways a literary text is 'scanned' in order to achieve an hermeneutic interpretation. In this context, the experience of the underlying structure as perceived is bound to the task of 'reading' the musical score, and the sonic internal representation of the musical piece is taken for granted and mainly confined to the realm of musical expertise. A Schenkerian analyst, therefore, learns how to build an interpretation of the musical piece. It follows from the above that the status attributed to structural hearing, or better, to the listener's awareness of the underlying voice-leading, takes for granted an aural component of the internal representation of prolongation, with the consequence of discouraging further explorations of its psychological plausibility. Because of this, the implications of the underlying structure, and of prolongation as experienced knowledge have not been thoroughly explored.

Second, and differently, according to Lerdahl and Jackendoff's *GTTM*, a listener is *ideally* competent (but competent anyway in the Chomskian sense) to *derive* a reductional

representation of a given musical surface. In this case, the cognitive approach will allow the performance of certain computational operations by means of which, in the end, an *experienced* representation of that musical surface will be obtained. Where is prolongation in this journey? Is this last prolongation the same prolongation that Schenker conceived of, except for the fact that when Schenker conceived prolongation, cognitive science had not yet emerged, and when it finally did, prolongation came to be conceived of as about computations with copies of pitch events that represent other events in the tree hierarchy?

The problem of the conflicting ontologies of prolongation, I think, is, in part, due to the conceptual models (Zbikowski, 2002) that dominate in the fields of musicology, music theory and music analysis on the one hand, and of cognitive psychology on the other. Is it, as Cook asserts, that the experience of prolongation is not about the way music *is* listened to but of how it *should be* listened to? (Cook, 1989). Or last but not least, is the only possibility for prolongation to have a place in cognition by reducing it to a series of computations, according to the constraints imposed by classic cognitive science?

It was hypothesized in this thesis that, over the course of this history, the consideration of the psychological status of prolongation lost an imaginative and metaphorical component that *is* implicit in the origin of this concept, emerging under the umbrella of Schenkerian theory.

Schenker managed to explain the understanding of the musical work based on an idealistic aesthetic conception, combined with a naturalistic view of music's unfolding as an organic force. He described the unfolding of the underlying musical structure as the manifestation of energetic aspects of the natural world. Implicit in this conception is a metaphorical component that maps aspects of reality with the understanding of music as organized sound (Lakoff and Johnson, 1999; Zbikowski, 2002, Saslaw, 1997-1998, Larson, 1997,

324

2004). The interplay of the dynamic forces that emerge from the unfolding of the musical form is interpreted as a course of action, which confers a sense of direction to music's deployment. Given that prolongation, as imagined sound, did not find a clear place in cognition, according to the terms of classic cognitive science, their psychological implications appeared unapproachable within the paradigms offered by that science and were largely unexplored.

Nevertheless, some evidence of the psychological status of prolongation had been already collected (Serafine et al. 1989; Martinez & Shifres, 1999b, c, 2000) and, according to these findings, prolongation, as perceptually experienced, deserved further investigation of its status in cognition.

It was assumed that the insights about the psychological plausibility that emerged from the prolongational theory were open to experimentation. From the point of view of classic cognitive science, there were aspects of the experience of prolongation that had not yet been explored, in spite of some indications that they might manifest themselves in music perception. The group of investigations that were pursued in the first experimental part of this thesis continued on lines partially opened up in previous research (see Chapter V). A hypothesis of constituency of prolongation was proposed, and its cognitive manifestation in music attending tasks was thoroughly explored.

While the experiments of Chapter V demonstrated that prolongation *is* a process that is cognitively relevant, they did not clarify the epistemological bases of such a process in cognition. As noted earlier, it was considered that the status of prolongation, in the terms proposed by Schenkerian theory, was germane to an idealistic component, investigation of which required an approach more aligned with alternative cognitive views of metaphorical thinking and imaginative listening than with those proposed within classic cognitive

science. Consequently, a viable hypothesis concerning the manifestation of prolongation as imagined knowledge was developed and pursued in order to be more precisely informative about the nature of its status in cognition. The experimental work of the second part of the thesis accounted for the idea of prolongation as a structural metaphor (see Chapter VII).

What follows is a summary of the experimental work that was conducted in this thesis.

VIII. 2 Summary of the experimental work

VIII. 2.1 The experience of prolongation as a constituent structure

In order to explore thoroughly the cognitive scope of prolongation, it was decided to investigate whether prolongational structure was manifest as a representation that could be used in the listener's cognitive activity with music. As mentioned above, previous research that found evidence of the manifestation and use of underlying representations in some cognitive tasks encouraged the further investigation of this concept. Following the experimental line that had been pursued before, the use of prolongation in other cognitive tasks was considered. This time the interest was placed in studying the moment-to-moment processing of the prolongational structure. Music attending appeared as an appropriate cognitive activity in which the use of prolongation could be investigated.

In order to test the experience of the prolongational structure as a tonal organization in which prolonging events are understood as *extensions* of a prolonged event, it was decided to hypothesize prolongation as having the status of a constituent structure. Investigating the ongoing experience of prolongation as a constituent organization should inform about the concomitant operation of a cognitive mechanism called mental retention (see Chapter I), in

which a tone is mentally represented and *kept in mind* until the next tone of a structural

upper voice becomes active.

To assign a constituent function to prolongation was to suggest that the linear continuity of

the unfolding voice-leading could be parsed into syntactical units with beginnings and

endings; it also meant that the constituent status of prolongational units could be

represented in terms of *boundary locations* (see Chapter V). On the other hand, applying a

hypothesis of constituency to prolongation meant that it might exhibit, to a certain extent, a

capacity to convey relative degrees of closure, as long as the underlying structure was

unfolded, and that this capacity could be a clue to identifying prolongational boundaries

throughout a particular voice-leading arrangement. That was the primary reason to

investigate the cognitive status of the constituent boundaries in terms of their relative

degrees of closure.

The aim of experiment V.1 was to determine precisely what was the influence that certain

aspects of the syntactic construction exerted on the experience of closure. The first section

of Chopin's Ballade No.1 in G minor was selected as the stimulus. The characteristic

motive repetition that the fragment presented was considered quite appropriate to the aim

of the study. Repetition would provide an opportunity to investigate closure, manipulating

similar musical materials and their syntactic functions, in different contexts of the

constituent structure of this musical passage.

Therefore, a further exploration of the perception of musical closure was accomplished,

with the aim of estimating whether the syntactical component of the first section of the

Chopin's Ballade had an effect on the perceived degrees of closure of the different

constituent units. Experimental conditions included the arrangement of the grammatical

context, according to both grammatically and ungrammatically arranged musical

sequences, in which concatenated and isolated units were presented to listeners. The experimental procedure included the use of a real time parsing paradigm.

The experimental results provided evidence of the sensitivity to closure during an ongoing listening situation. In particular, in the case of embedded contexts, these results illuminated the capacity of the context to confer conditions of relative coherence on the grammatical component of a musical passage, and to elicit expectations about the ongoing syntactic course of that passage. Moreover, they highlighted the idea that contextual coherence helps keep in memory the function of a syntactical unit every time the syntactical order is preserved. It could be said that contextual coherence provides an opportunity to establish mental connections between separated events. Therefore, the cognition of closure appears as a generic psychological process in the perception of tonal music.

Central to the purposes of the present thesis was the need to obtain behavioural evidence of the cognitive status of *constituent boundaries* in grammatical contexts. While showing a capacity to account for differences in the internal articulation of constituent structures, closure validated the cognitive status of boundary locations as focal points in the processing of a grammatical component of tonal music. If boundaries had cognitive reality in constituent organizations, then it was possible to think that they could be used to test the cognitive experience of *prolongational* contexts.

This hypothesis was tested in experiment V.2. A music attending task provided the opportunity to investigate an assumption of prolongation as a constituent organization. Prolongational structures were understood in terms of dependency relations that took place inside the constituent, establishing differences between pitch events in terms of their relative structural importance. It was assumed that within a musical passage, the listener would experience certain events as being related to, and dependent on, some other non-

immediate events. Sensitivity to this structural organization should elicit a kind of mental operation in which two events were experienced simultaneously but at different hierarchical levels: the head of the constituent would be experienced as the dominant event and the prolonging pitch events would be experienced as attached to the head and also as eliciting expectations of continuation toward the following constituent head.

Given the evidence of the cognitive status of constituent boundaries found in experiment V.1, they were employed as musical markers to test a hypothesis of constituency of prolongation. If mental retention of the constituent head all through the prolongational unit occurred, the listener's expectations of continuation elicited by the prolongational extension of that structural event would, in fact, end at the boundary of that unit. If, by means of dependency, the structural note of the prolongational unit was extended and mentally retained all along the unit until the boundary was reached, a different attentional response was expected at-boundary locations, than the response expected at within-constituent locations. Thus, it became necessary to monitor the attending process at these two different locations and to find a measure that was useful to provide account of this issue. A subject's reaction time response appeared as appropriate to bring the solution that was needed.

Constituent units that presented instances of close prolongations were selected as stimuli. They were treated by placing a click (an extraneous signal used in linguistic studies) in two different positions, according to whether it occurred at a boundary location or out of a boundary location. The experiment used a divided attention task, with musical listening as the primary task and click-detection as the secondary task. Participants were required to press a key as soon as they detected the click. It was predicted that RTs would be faster for clicks located at prolongational boundaries and slower for clicks located before such

boundaries, on the assumption that, according to what theories of information processing posit, attention would be maximal with a minimum of cognitive processing at the boundary location. It was assumed that the degree of closure at the boundary location would result in a sense of fulfilment of the prolongational extension and, as a consequence, in decay of the cognitive effort, compared to what happened within the prolongational unit, where the head of the constituent was being mentally retained in the moment to moment processing of prolongation.

The experimental results found behavioural evidence of the listener's sensitivity to prolongation during music attending. The click-detection technique proved to work well with prolongation. SRTs were significantly different according to whether clicks were located within the prolongational units or at their boundaries. Thus, the psychological validity of the prolongational boundary as a focal point in music attending informed a status of *prolongation as a constituent percept*.

As noted in Chapter I, theories based on the concept of a computational mind highlight the importance of the structural limits of the constituent units in organizing the listener's parsing activity. Some locate structural priority in the organization of the rhythmic-metric component of temporal information (*GTTM*); others focus on the effect of closure features (Narmour's *I-R*). Evidence of a metrical effect in RTs was not found. Contradicting theoretical assumptions based on the structural priority of rhythmic-metrical factors in the building of reductional representations, the collected evidence indicated that listeners seemed to focus more on the continuity of linear aspects of the pitch organization, when they had to grasp the meaning of prolongational units during the moment to moment attention to music. According to our results, it seems that it is pitch more than rhythm that is the component that *has* structural priority at the prolongational boundary area.

The additional experimental evidence that was obtained in experiment V.3 enhanced the scope of the study of prolongation as experienced, using open prolongations, that is, structures that contain suffixes and/or preffixes in their external limits. Similar results to those found in experiment V.2 were obtained. The use of open prolongations provided evidence of the participants' sensitivity to left and/or right prolongational branchings in reductional representations.

Experiment V.4's main purpose was to explore the participants'sensitivity to prolongational structure, testing the same cognitive process (attention to music) but using a click location task. It was expected, as in previous investigations on speech processing, to reveal a tendency to migrate the location of clicks to boundaries between groups. Results did not seem to support the migration effect in the experience of prolongational structures. Except for a small tendency to migrate clicks backwards to the boundary region of the prolongational units, the click-location activity did not exhibit a clear migratory effect in the location responses.

The click-location paradigm did not appear as appropriate to study the prolongational component in constituent structures. Compelled by the expectation of the occurrence of a single event (the click) throughout the flow of musical events, in order to provide a post hoc response about its location, the participant appears to activate a parsing strategy based on a clock-like beat division of the temporal continuity of the voice-leading unfolding. This outcome is interpreted as a plausible explanation for the lack of accuracy in the participants' responses that, by the way, is similar to the uncertainty found in previous experiments that used this experimental technique (see Chapter IV).

In summary, the findings of experiment V.1 to V.4 inform music cognition research in the following ways:

(i)   The constituent status of closed and open prolongational structures has been

         demonstrated for music attending tasks at foreground reductional levels of the

         underlying hierarchy of tonal music. The evidence suggests that the boundary area

         of prolongational structures *is* a focal point with clear status in cognition.

(ii)   The thorough exploration of different music attending methodologies

         accomplished in the current studies indicated that the click detection and not the

         click location technique is a suitable procedure in accounting for the on-line

         listening experience of prolongation. These differences in suitability can be

         attributable to the distinct types of temporal experience of the music unfolding that

         they activate. Attention to music, while prompting expectations about incoming

         information, seems to be related to time estimation, that is, to the way temporal

         intervals are phenomenally filled-in. However, according to Epstein (1995), there

         is a duality in the sense of time in the experience of music: on the one hand, it

         involves the activation of a chronometric component, that is, a clocklike temporal

         sense that elicits a beat-compatible division of the temporal continuity into discrete

         pieces of musical information; on the other hand, the sense of time seems to be

         more akin to the intrinsic configuration of the musical piece, activating a sense of

         time that emerges from the individual's internal experience of the piece as a whole.

         Although these temporal dimensions shape musical time differently, both are

         assumed to be available during music listening. At the core of prolongational

         theory is the idea that the temporal unfolding of the underlying structure depicts a

         dynamic deployment that is evidenced in the voice-leading arrangement. This view

         of the dynamic quality of the underlying unfolding is related to the second

         dimension of the sense of time. The results of experiments V.2 and V.4 show

clearly that the click-detection task does not interfere with the on-line sense of temporal continuity that fits better with the experience of the voice-leading; in this ongoing process, the boundary of a prolongational unit appears naturally to perception as a focal point at which a relative sense of completion is reached. RTs accounted for such cognitive behaviour. On the contrary, the click-location procedure seems to prime a sense of temporality that is not congruent with the experience of the underlying structure in the terms proposed by prolongational theory.

(iii)   The absence of differences between musicians and non-musicians in their sensitivity to prolongational structures confirms previous findings about commonalities in the accomplishment of groups with different degrees of musical expertise when they are assessed at a basic level of cognitive achievement. According to these results, the cognitive status of prolongational structure counts as "basic knowledge", at least at a foreground reductional level of the underlying hierarchy.

(iv)   In the end, an implied ontology of the constituent function of prolongational structures at foreground reductional levels of the underlying hierarchy of tonal music emerged from the experimental work conducted in this thesis.

VIII.2.2 The experience of prolongation as a structural metaphor

Although the previous experiments had demonstrated that prolongation was cognitively relevant, they had still not clarified the epistemological bases of such processes in cognition. According to Schenkerian theory, prolongation appeared to be more than just a

constituent percept. In Schenker's seminal proposal of the underlying structure, there was an imaginative component that the idea of constituency did not fully explain.

Prolongation fundamentally concerns entities that, in part, have no physical record. The description of prolongation in terms of a feature of the voice-leading unfolding, where a tone is extended all through a musical passage and *remains active* in the musical flow of events, even though it has no substance, might be viewed more as metaphorical rather than as formal and literal. The concept of mental retention as a process that is involved in the experience of prolongation might very well be related to an imaginative component in music cognition. On the other hand, as noted earlier, the aesthetic principles that are germane to Schenker's view of the underlying musical structure reflected, in part, the Kantian idea of the understanding of form as an *a prioristic* configuration, of which an image is elicited by the concrete work of art.

Schenker conceived the notion of a piece of music based on the metaphor of a live organism, that is to say, as an entity that is governed by some biological principles emerging from the natural world. The concept of prolongation that he developed is intrinsically embedded in this ontology. In this context, the unfolding of the underlying structure by means of prolongation implies the elicitation of several principles that metaphorically invoke forces in movement that struggle in fulfilling a tonal goal. The idea of tonal tension, understood as a course of action that takes place all through the unfolding of the fundamental structure by means of the voice-leading procedure, for Schenker has a dramatic connotation (Schachter, 1999).

The psychological component in Schenker's view of the underlying structure, if there is one, emerges from the application of the above noted principles to the experience of the

musical sounds. He was the first to propose what Cook ably asserts (Cook, 1989) – that what is similar on the surface is experienced as different according to the context.

The language used by Schenker to describe the unfolding of the underlying organization of musical pieces includes linguistic metaphors such as traversal of a path, directed motion, different forces in operation, and achievement of a goal, among others (see Chapter VI); these metaphorical concepts account for a view of the underlying structure in idealistic terms. The question is: of what kind are these metaphors that emerge from the Schenkerian argument about the prolongational structure of music?

The imaginative component that is intrinsically embedded in Schenker's conception of the underlying structure was investigated in the second experimental section of this thesis (Chapter VII). It was realized that the question about the imaginative status of prolongation could not be properly answered in the terms proposed by traditional cognitive science, and that its investigation needed to be founded in alternative views of human cognition related to the study of embodied knowledge and metaphorical thinking.

It was necessary to elaborate a hypothesis of the experience of prolongation as imagined cognition. Accordingly, it was hypothesized that prolongational structure would be experienced in terms of a *structural metaphor* (Chapter VI). A theoretical framework that had recently been developed in the fields of cognitive science and cognitive linguistics about the identification of a metaphorical component operating in human cognition, appeared to be the proper context in which to locate this hypothesis. According to the main findings in this field, our general understanding of reality is in part a consequence of the embodied structures or forms of imagination that grow out of our bodily experience. Imagination plays a central role in all that concerns the processes of understanding, reasoning and attribution of meaning to our everyday experience. The image-schematic

335

structures that organize our basic knowledge are acquired during the course of our sensory motor activity and are activated unconsciously when we are grasping meaning in different areas of knowledge.

As noted earlier, some of the theoretical postulates about the underlying musical structure had already been elaborated as conceptual metaphors (see, for example, Zbikowski, 2002; Saslaw, 1996; 1997-1998; Larson, 2004). These descriptions use metaphorical language to analyse the behaviour of structural features in shaping the dynamic organization of tonal tension in musical pieces. However, the use of image-schematic knowledge in the listener's experience of the prolongational structure had not been tested experimentally so far. On the assumption that metaphorical thinking, present in the language used to conceptualise music, might to some extent, model the understanding of structural aspects of music, the exploration of the cognitive scope of these principles as applied to music *as experienced* was accomplished.

It was hypothesised that the temporal unfolding of the underlying structure elicits a kind of imaginative listening, by means of which a sense of direction is conveyed to the musical flux, and that some of the linear techniques which produce the melodic and harmonic prolongations, interacting with cognition, would prompt the activation in music perception of an imaginative and/or creative process, by means of which it is possible to hear *x as y* (Larson, 1997) that it to say, to assign sounds to categories. The operation of those tension-based features of the pitch event's organization (such as the tendency of the fundamental line to go downward, the tendency of an unstable tone to move towards the closest tone at different levels of the underlying hierarchy, and the general tendency of the musical discourse to reach the tonal goal) should convey the kind of dynamic quality to music unfolding that we were interested in investigating.

336

It was proposed earlier (see Chapter VI) that the application of the concept of forces in movement to the analysis of music linked everyday life notions of mechanics, such as the concepts of generation, flux and transformation of energy, and used them to attribute an energetic quality to music unfolding. It was assumed in this thesis that this energetic quality that emerges from the unfolding of the sonic organization of the musical piece is mapped onto the dynamic patterns that emerge from the listener's embodied experience of forces in movement. According to our hypothesis, this is the meaning conveyed by Schenker to the concept of organic coherence.

The interest in exploring a more comprehensive dimension of the prolongational structure encouraged the search for a structural configuration that might account for such a level of representation. The *interrupted structure*, one of the most important Schenkerian descriptions of musical form in his theory of the underlying structure, was selected as the object of experimental testing.

A cognitive view of this structural principle was proposed in terms of the conceptual metaphor of *musical forces.* In experiment VII.1 it was hypothesised that the interrupted structure would be metaphorically experienced in terms of the *blockage-release of blockage* image-schema, a sub schema of the force image-schema (Johnson, 1987). Once that image-schema is activated, by means of a cross-domain mapping process, it projects itself onto the underlying organization of the tonal piece, priming a categorization process by means of which the listener *understands* the interrupted structure as a sonic unfolding of such force schema.

In the experiment that was conducted, it was expected that participants brought about a cognitive process where they understood the interrupted structure in metaphorical terms. Each musical fragment was contrasted with different musical reductions (understood as

simplifications of the surface level of the piece) of the fragment. A priming paradigm was used to activate the image-schema that the participant was assumed to use in the process of cross-domain mapping, to hear A in terms of B, A being the musical fragment and B the structural features highlighted in the musical reductions. During the experimental procedure, after being primed with the visual animations, participants listened to the musical fragments and their correspondent reductions and estimated the goodness of fit between them. The key assumption was that the visual priming that was assumed to activate the blockage-release of blockage image-schema would elicit in the listener a tendency *to hear the musical piece more in terms of the interrupted structure than of other features*. Therefore, the participants would tend to match the reduction that conveyed the structural feature intended to correspond to the image-schema that was activated by the visual priming to a greater extent. When the force image-schema was not activated (because a different visual priming was used) the matching between the voice leading reduction and the musical fragment would be lower.

The results supported the hypothesis. Participants understood the interrupted structure of tonal compositions in metaphorical terms. It was concluded that metaphorical thinking contributes to shape the experience of musical structure. Once a force image-schema has been activated, a cross-domain mapping process takes place and the subject uses the image-schematic knowledge of force relationships in order to hear the underlying voice-leading of the interrupted structure (highlighted in the musical reduction) in terms of purposeful movement. Therefore, the evidence indicates that the interrupted structure is *metaphorically experienced* when the listener is capable of understanding the harmonic-melodic process of tension that unfolds the relationship 3-2//3-2-1 in the upper voice and I-V//I-V-I in the bass, as purposeful movement that is blocked and subsequently released, to

finally arrive at the resolution of the tension that was interrupted in the middle of the formal phrase division.

Two problems that were identified in the research about the internal representation of musical structure I have tried to overcome in this thesis. The first of them emerges when the study of the listener's representation of music in the Western culture is approached, and is related to an implicit use of the musical score *as* a reification of the real music in the experimental procedure (Cook, 1989; 1990). It infiltrates the analysis and modelling of music cognition to the extent that it is not clear which is the implicit role that the musical score *as the* representation of the musical structure plays in psychological theories of music apprehension. It is also, in part, a criterion used to establish differences between sophisticated listeners and everyday listeners, and their correspondent grouping in populations of musicians and non-musicians, that are so pervasive in experimental studies. Moreover, it is embedded in the methodological procedure (Rosner, 1984; Cook, 1989): in many cases, experiments that were intended to test even the simplest musical representations make use of scores and/or feature-like-scores diagrams in order to *activate* this reified musical representation (see Chapter IV); therefore, it is common that conclusions applied to music listening that are derived from the results of such experimental work are indeed more the result of a listening activity that makes use of visual score-like representations of the music.

This problem has been identified by Cook (1989) for example, in Deliège's (1987) experiments on aspects of Lerdahl and Jackendoff's *GTTM*, in which she used score-like visual notations to account for the listener's internal representation of grouping mechanisms. It is not found in the experiments conducted in this thesis because, in order to test the participant's representation of the prolongational structure, musical scores were not

used. What I was interested in was the internal representation of some of the postulates about the prolongational structure *without making use of the mediation of the musical score*. I wanted to explore the cognitive reality of the representation of prolongation *as experienced through listening* and not as contemplated in the musical score, as analysts do.

The second problem that has been identified is related to the length of the samples used in experiments that test the cognitive representation of the tonal structure. It has been proposed that conclusions about such representation are frequently derived from the use of very short fragments that in fact cannot be informative of the cognitive processing of larger-scale musical passages. Although the samples used in my experiments, in many cases were no longer than a complete phrase, some examples in the second experimental part used whole microforms, and all the musical samples in experiment VII.1 were much longer than the psychological present (see, for example, figures 7.6.a, 7.7.a, and 7.9.a, in Chapter VII), which also answers Rosner and Meyer (1986), and Cook's (1987; 1989) own concerns about the limited scope of experimental tests of the listener's awareness of musical structure that avoid the experience of larger-scale underlying representations.

In summary the findings of experiment VII.1 inform music cognition research in the following ways:

(i)　　An imaginative component of the prolongational structure was investigated and it was found that the underlying configuration of a musical composition, interacting with cognition, prompts the activation in music perception of an imaginative process, by means of which it is possible to categorize musical sounds in terms of structural features.

(ii)   Results demonstrate that the interrupted structure is *metaphorically experienced* as purposeful movement. By means of a cross-domain mapping process, the meaning of this structural feature can be understood as a course of action that, while pursuing the final goal, is blocked and subsequently released, to finally arrive at the resolution of the tension at the end of the phrase.

(iii)   The epistemological base of prolongation in cognition was explored and it was found that it has a relevant status as imagined cognition. It was demonstrated that the force metaphor is more than just a linguistic construction used in the theoretical discourse about music. It appears to operate as an *idealized model of cognitive processing* that listeners activate during their experience of the underlying musical structure (Lakoff, [1987] - 1990).

VIII. 3 CONCLUSION

This thesis investigated the psychological status of a concept that is part of music analytical discourse. Is the meaning of the piece of music defined by the perception of musical sound? Is it defined according to what analytical descriptions suggest? Is it the product of the physical actions involved in performance? Or, last but not least, is it the result of performing mappings between the musical domain and other domains of human experience? These are questions that have been answered differently according to the conceptual and cultural constraints that have been and are still prevalent in the West for more than a century.

It was considered here that the prolongational structure, developed for the first time in Schenkerian theory, had a descriptive value that needed to be explored in terms of its cognitive status as experienced knowledge. If theories pertaining to the principles of musical structure develop assumptions about certain phenomena of music cognition, then it is the work of the psychology of music to derive formulations that turn out to be demonstrable by means of experimentation. As a result of the experimental work that was conducted here, it was found that, at foreground reductional levels close to the musical surface, encompassed by a temporal span that fits the psychological present, it may be possible to characterize the experience of the prolongational structure as a constituent percept. Moreover, it was found that at more comprehensive and deeper levels of the underlying hierarchy, the experience of prolongational structure seems to be constrained by the operation of a kind of imaginative representation, metaphorical in nature. This kind of mental representation fits better with Schenker's idea of the underlying structure in music.

Schenker was a pioneer in proposing a grammatical component in the analysis of the underlying structure. Indeed, the most striking connection between his theory of hierarchical levels and the theory of generative grammar is the shared belief that a coherent and unifying set of principles underlies creative human behaviour (Keiler, 1978, p 174). In this thesis, a syntactical component of the prolongational structure was investigated according to the principles of constituency and dependency, and it was demonstrated that the moment to moment dynamic process of attention to the prolongational structure is cognitively relevant in the terms proposed by the theory of musical syntax. However, these experiments provided a necessary but not sufficient clarification of the epistemological bases of such processes in cognition. There were questions that had not been answered: they were related to the status of the metaphorical extension over which the process of mental retention seemed to operate (Benjamin, 1982).

In the second experimental part of this thesis, another syntactic attribute of the underlying hierarchy was tested. But this time the focus of the inquiry was in the epistemological status of prolongation as imagined sound in cognition. Following this concern, the study of the interrupted structure was approached from the point of view of the theory of embodied knowledge and metaphorical thinking. Thus, the interest was not in just exploring the status of grammaticality of the prolongational structure, in as much as in seeing if the deployment of such underlying component had a capacity to convey to cognition a dynamic representation as purposeful movement. The ultimate purpose of this thesis was that of exploring whether prolongation, as a feature of underlying voice-leading, contributed to preserve aspects of the idea of good form (understood as the *Urform* in Goethe, or its parallel *Der freie Satz*, in Schenker (Keiler, 1978) as manifested in cognition.

It was stated earlier that the relationship between music theory and psychology of music was constrained by the conceptual models that were prevalent in both disciplines, and that some epistemological concerns emerged when the psychological validity of a musicological concept was investigated. The complexity of the relationships between the make up of a musical piece, its perceptual experience and the validity of a musical theory in accounting for such perceptual experience has already been acknowledged (Cook, 1989; Brown and Dempster, 1989; Boretz, 1989). For example, it is true that there is no obligation for a musical analysis to describe how people really listen to music; in any case, a musical analysis can develop a potentiality to describe how music 'should' be heard. According to Cook (1987; 1989), that is the sense of Schenker's theoretical proposal. But even if it is true that a musical analysis is a way of describing a 'reading' of the musical piece, which has the capacity to build a bridge between the composition and the perceptual mechanisms of the listener, the question of the extent to which a musical theory somehow reflects the auditory experience of music still remains.

Schenker's conception of the underlying structure, as we saw, is intrinsically related to his music analytical practice. Although it is impossible to say to what extent the musical intuitions of the analyst may or may not be assimilable to the musical intuitions of the listener, one possible way to focus this problem is conducting experimental research. However, according to Cook (1989) in spite of the efforts of experimental psychology to explore the perceptual status of music theoretical developments, the methods chosen in the analytical tradition - including Schenker's own conception of the underlying structure - might have little relation to the ways music is really heard. Thus, the value of a Schenkerian analytical interpretation would lie, not in proving that the theory is correct, but in presenting a convincing reading of the structure of the piece, an explanation that can

be susceptible to intersubjective acceptability. In his words: "...we do not accept an analytical interpretation just because we can "hear"it. [...] We accept it because we find it persuasive" (p.135). That is to say, what matters here is the quality or demonstrative value of an analysis in providing a convincing picture of the musical piece.

However, the heuristic procedure that is followed when an analysis of the underlying voice-leading is performed might collapse with the concept of "cognitive reality" according to the constraints posed by cognitive psychology. Because, as suggested by Nattiez (1992) "...a musical analysis never proposes a realistic image of the work, which reveals its essence, but a *construction* which selects in the object of study a number of features considered as appropriate to give an account of it. The problem consists of understanding the principles that are the basis of the selection of those features." (p. 537). According to this, a 'correct' understanding of the principles of voice-leading would result in an appropriate interpretation. But this interpretation would be legitimate if it has the capacity to provide sense or to bring a coherent explanation of inter-subjective value. In parallel to this Imberty (1998) proposes:

"It is apparent that [an interpretation] could never be proved because it is not possible to prove the truth of an interpretation. Nevertheless, there are some interpretations that are more legitimate than others, or that are worth being valued because they involve both the interpreter and the addresser of the interpretative discourse. But, in order to make this manoeuvre consistent, it is necessary for the interpretation not to be wholly incompatible with the *real facts* and it is also necessary to circumscribe it to a specific context. Once more, this will not allow saying that the interpretation is scientific: the same fact may be interpreted in different ways" (p.105, my italics).

According to Imberty, even if it is true that it is not possible to prove the truth of an interpretation when an analytical method, such as a Schenkerian analysis is applied, there are, nevertheless, some interpretations that are more reliable than others. In order for an interpretation to be reliable, it should be acknowledged an extended value. This quality of comprehensiveness is achieved only if the interpretation *is prepared*, that is to say, if the categories of analysis are carefully considered and *enough data* are collected.

In order to consider the nature of the data relative to the analysis of prolongation as experienced, some of the sources revised here circumscribe the concept of cognitive reality within the framework of Schenkerian reductional theory: they assert that such *reality* seemed to arise from the description of an "almost-psychic" domain, something that in fact is not real, that indeed is a hermeneutic. That is to say, they emerge from a reality that might be assigned a status of a physical fact; which could be transformed into an object, which could be *reified* (Cook, 1990).

The weight of the data resulting from the investigation conducted in this thesis indicate that the reality of prolongation as experienced in cognition can be understood as emerging from the activation of an *idealized cognitive model* (Lakoff, 1987) which, according to the theory of metaphorical thinking, maps physical 'facts' onto musical 'sounds', by means of a cross-domain mapping process, in which image-schematic aspects of general knowledge are used to produce metaphorical interpretations of musical sound.

An objective formulation and explanation of musical experience could be based on a highly deterministic theory that explains the individual's experience in terms of what are assumed to be real causes. But it could be argued, instead, that the so-called causes are not in reality causes at all, because they have no existence outside the act of musical experience; they are explanatory metaphors or fictions, which validity should not be denied

346

(Zbikowski, 2002). They are, instead, qualified as social constructions that belong to the musical culture, in the sense that they depend for its very existence on the availability of inter-subjective representations for music: "Because a musical culture is in essence no more and no less than a body of knowledge shared between culture-members" (Cook, 1990; p.241-242). Following this idea, the cognitive assumptions of theoreticians can be transformed into hypotheses subjected to experimental testing, allowing the formalization of those aspects of the theories that have not yet been formalized. This has been the main goal of this thesis.

The use of metaphor as a feature of musical understanding has already been acknowledged, at least from a theoretical point of view. As Zbikowski (1998; 1997-1998), Cook (1990) and Scruton (1983) propose, music analysis and, by implication, musical experience - reinforced now by the results of this thesis - are essentially metaphorical. Making a distinction between the material factuality of sound and the intentionality of music as a mental construction, Cook says: "A Schenkerian analysis is not a scientific explanation but a metaphorical one; it is not an account of how people actually hear pieces of music but a way of imagining them" (1990, p 4). According to Zbikowski (1998) the metaphoricity of musical understanding is inescapable and connected with the establishment and development of music cultural concepts. What we experience is what we know about music and depends on the images by means of which we represent the commonalities and the differences from within a musical culture. That is what gives a musical culture its identity.

The recent body of research in cognitive science, embodied cognition and cognitive linguistics supports this assertion (see Chapter VI). It is not that the characterization of musical features is the same as the characterization of physical features, as Zbikowski

(1998) proposes. Thus, for example, physical motion is not the same as musical motion. The latter is mapped from the physical to the musical by means of correlations between both domains of experience. This is what metaphorical transference is about. The theory of image-schemas provides an alternative explanation to human cognition, in presenting a view of understanding that is not the result of an after the fact reflection on prior experience but a way of *understanding while experience is taking place*. So, what we understand about music is what we experience of it, and this is deeply rooted in our culture, history and bodily mechanisms that blend together, in order to form those image-schematic structures with which we operate to assign meaning to music.

If music is essentially a kind of activity that, in words of Cross (2005) 'embodies, synchronizes and intentionalizes sound in action', the study of those aspects of the embodied experience that activate imagination should be of high potential value to pursue future investigation on the nature of music as experienced.

I proposed in this thesis that there is an imaginative component in Schenker's conception of the underlying musical structure that those theoretical models that develop psychological assumptions based on the classic cognitive science approach did not appear to capture fully. The concept of tension/relaxation as approached in these theoretical formulations does not explain the nature of prolongational structure, because the kind of abstraction that the tension/relaxation concept proposes does not properly match the ontology of prolongation as formulated in Schenkerian theory. Generative theory alone fails in explaining the content of what takes place in the course between two structural events, that is to say, between an abstract representative and another abstract representative, because it concerns itself with selected events that are extracted from the sequence of events and transferred to a higher level of representation. On the other hand,

metaphorical extension, as proposed in Schenkerian representations, is a kind of abstraction that accounts for what takes place in the *path* between events that form part of the structure that contains them. A metaphorical extension fits better with a characterization of the underlying structure as an inclusional hierarchy of events.

For that reason, the theory of image-schemas, that deals with abstract representations of dynamic relations between the components of a given knowledge domain, is more appropriate to account for the Schenkerian's conception of the underlying structure as an archetypal synthesis of a fundamental tension between opposites (I and V) and also - and mainly - of a path to traverse them (the voice-leading that unfolds the fundamental line). The generative hierarchy with its derivative packets fails to account fully for this comprehensive phenomenon. There is a vacuum between one event and the event that follows in those information packets. On the contrary, by means of operation of imaginative cognition, those packets are filled in with dynamic content, with force, with sense of direction, with blockages and returns. As said before, the time-space acquires a dramatic dimension.

The aim of this thesis was in a sense similar to the aim of an analytical interpretation. It was not intended to demonstrate the truth of Schenkerian theory, but to enquire the psychological plausibility of some of the ideas captured by Schenker the analyst, that is to say, to see if they had a place in the heard cognition. This is not to say that the cognitive experience of listening to the musical structure is similar to the cognitive experience of analysing the musical structure. As noted earlier (see Chapter II) some ideas that are born as analytical intuitions can afterwards acquire cognitive status if research is pursued so as to endow them with psychological validity. According to this argument, music theorists' intuitions about the ways in which people experience music are not scientific

generalizations, but nevertheless, they are the result of their own and deep musical craftsmanship. Analytical statements seem to arise from the interplay between analysts' own hermeneutics and the knowledge coming from their conscious awareness of their own perceptual processes. Such statements might be ascribed to the domain of so-called *folk psychology* (Cross, 1998) as long as they account for both analysts' intuitions and the particular ways in which they use their musical perception to contrast their own acts of analysis. This view somehow confronts the view of music psychology, which, according to Cook (1990), focuses on the study of the *musicological perception*, implying that it tackles simple music-theoretical notions, to an extent that undermines its capacity to explain *musical perception*, which is the result of the everyday musical experience.

However, the results of the research reported in this thesis, applying the theory of embodied knowledge to the study of the prolongational structure, bring alternative answers to these apparently contrasting views of the understanding of music as perceived. They provide indicators that attest to the value of our embodied everyday experience in the understanding of music, even when applied to listening situations that could be characterized as instances of musicological perception. They also suggest us that we use embodied basic knowledge to account for a complex and abstract phenomenon such as the grasping of the underlying musical structure. And, finally, they indicate that it is not necessary that psychology of music deals only with extremely simple music theoretic notions in order to explain the nature of musical experience. It can enhance the scope of its work in as much as it also enhances the view of what cognition is about, taking into account other aspects of human experience such as those conveyed by the theory of embodied knowledge.

INDEX OF FIGURES AND TABLES

REFERENCES

Agawu, V. K. (1989). Schenkerian notation in theory and practice. *Music Analysis,* 8(3), 275-301.

Agawu, K. (2004). How we gott out of analysis, and how to get back in again. *Music Analysis,* 23(ii-iii), 267-286.

Aiello, R. and Sloboda, J. A. (eds.), (1994). *Musical Perceptions*. New York: Oxford University Press.

Arnheim, R. (1984). Perceptual dynamics in musical expression, in *New Essays on the Psychology of Art.* Berkeley: University of California Press. pp. 214-227.

Askness, H. (1997). A cognitive approach in musical analysis: metaphorical projection in music, in A. Gabrielsson (ed.), *Proceedings of the Third ESCOM Conference.* Uppsala, Sweden: Upssala University, 551-556.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 11(4), 629-654.

Barsalou, L. W. (1989). The instability of graded structure; implications of the nature of concepts, in U. Neisser (ed.), *Concepts and Conceptual Development.* Massachusetts: Cambridge University Press, 101-140.

Beach, D. (1983). Schenker's theories: a pedagogical view, in D. Beach (ed.), *Aspects of Schenkerian Theory.* New Haven: Yale University Press, 1-37.

Beach, D. (1984). Motive and structure in the Andante movement of Mozart's Piano Sonata K. 545. *Music Analysis,* 3(3), 227-241.

References

Benjamin, W. (1981). Pitch-class counterpoint in tonal music, in R. Browne (ed.), *Topics in Music Theory*. New York: Academic Press, 1-32.

Benjamin, W. (1982). Models of underlying tonal structure: how can they be abstract, and how should they be abstract. *Music Theory Spectrum*, 4, 28-50.

Berent, I. and Perfetti, Ch. A. (1993). An on-line method in studying music parsing. *Cognition,* 46, 203-222.

Berry, Wallace ([1966] - 1986) *Form in Music*. Second Edition. New Jersey. Prentice-Hall.

Bharucha, J. J. (1984a). Anchoring effects in music: the resolution of dissonance. *Cognitive Psychology,* 16, 485-518.

Bharucha, J. J. (1984b). Event hierarchies, tonal hierarchies, and assimilation: a Replay to Deutsch and Dowling. *Journal of Experimental Psychology, General,* 113(3), 421-425.

Bharucha, J. J. (1996). Melodic anchoring. *Music Perception,* 13(3), 383-400.

Bharucha, J. J. and Krumhansl, C. L. (1983). The representation of harmonic structure in music hierarchies of stability as a function of context. *Cognition,* 13, 63-102.

Bigand, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music,* 18, 45-59.

Bigand, E. (1994). Contributions de la musique aux recherches sur la cognition auditive humain, in S. Mc Adams and E. Bigand (eds.), *Penser les Sons. Psychologie Cognitive de l'Audition*. Paris: Presses Universitaires de France, 249-298.

Blasius, L.D. (1996). *Schenker's Argument and the Claims of Music Theory*. Cambridge: Cambridge University Press.

Boltz, M. (1989). Perceiving the end: effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance,* 15(4), 749-761.

Boltz, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics,* 53(6), 585-600.

Boretz, B. (1989). The logic of what? *Journal of Music Theory,* 33(1), 107.

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

Brower, C; (2000). A cognitive theory of musical meaning, *Journal of Music Theory*, 44(2), 323-379.

Brown, M. and Dempster, D. (1989). The scientific image of music theory. *Journal of Music Theory,* 33(1), 65-106.

Brown, S. W. and Boltz, M.G. (2002). Attentional processes in time perception: effects of mental workload and event structure. *Journal of Experimental Psychology: Human Perception and Performance,* 28(3), 600-615.

Burkhart, Ch. (1978). Schenker's motivic parallelisms. *Journal of Music Theory,* 22(2), 145-176.

Cadwallader, A. (1983). Motivic unity and integration of structural levels in Brahms's B minor Intermezzo, Op. 119, No. 1. *Theory and Practice,* 8(2), 5-24.

Cadwallader, A. (1988). Foreground motivic ambiguity: its clarification at middleground levels in selected late piano pieces of Johannes Brahms. *Music Analysis*, 7(1), 59-91.

Cadwallader, A. and Gagné, D. (1998). *Analysis of Tonal Music: a Schenkerian Approach.* New York: Oxford University Press.

359

Cadwallader, A. and Pastille, W. (1992). Schenker's high-level motives. *Journal of Music Theory,* 36(1), 117-148.

Cambouropoulos, E. (1997). The role of similarity in categorisation: Music as a case study, in A. Gabrielsson (ed.), *Proceedings of the Third Triennial ESCOM Conference,* Uppsala, Sweden, 533-538.

Caplin, William E. (1998). Classical Form. A Theory of Formal Functions for Instrumental Music of Haydn, Mozart and Beethoven. Oxford: University Press.

Carterette, E. C. and Kendall, R. A. (1999). Comparative music perception and cognition, in D. Deutsch (ed.), *The Psychology of Music.* San Diego: Academic Press. pp. 725-791.

Chang, J. Feldman and Narayanan, S. (2004). Structured connectionist models of language, cognition, and action, in *Proceedings of the Ninth Neural Computation and Psychology Workshop*. Plymouth: University of Plymouth. pp. 57-68.

Cohen, G. (2000). Hierarchical models in cognition: do they have psychological reality? *European Journal of Cognitive Psychology*, 12(1), 1-36.

Cohn, R. and Dempster, D. (1992). Hierarchical unity, plural unities: toward a reconciliation, in K. Bergereon and P. Bohlman (eds.), *Disciplining Music. Musicology and its Canons*. Chicago and London: The University of Chicago Press, Chapter 9, 156-181.

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception*, 5(2), 197-206.

Cook, N. (1989). Music Theory and good comparison: a Viennese perspective. *Journal of Music Theory*, 33, 117-141.

Cook, N. (1990). *Music, Imagination and Culture*. Oxford: Oxford University Press.

Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum,* 23(2), 170-195.

Cook, N. (2002). Epistemologies of music theory, in T. Christensen (ed.), *The Cambridge History of Western Music Theory*. Cambridge. Cambridge University Press, Chapter 3, 78-105.

Covington, K. and Lord, Ch. H. (1994). Epistemology and procedure in aural training: in search of a unification of music cognitive theory with its application. *Music Theory Spectrum*, 16(2), 159-170.

Cross, I. (1998). Music analysis and music perception. *Music Analysis,* 17(i), 3-20.

Cross, I. (2003). Music and biocultural evolution, in M. Clayton; T. Herbeert and R. Middleton (eds.), *The Cultural Study of Music*. New York: Routledge.

Cross, I. (2005). Music, meaning, ambiguity and evolution, in D. Miell, R. MacDonald and D. Hargreaves (eds.), *Musical Communication*. Oxford: Oxford University Press. pp. 27-44.

Cumming, N. (1991). Analogy in Leonard B. Meyer's theory of musical meaning, in J. C. Kassler (ed.), *Metaphor: A Musical Dimension.* Sydney: Currency Press.

Damasio, A. (1994). *Descartes Error.* New York: Quill.

Damasio, A. (1999). *The Feeling of What Happens.* London: Vintage.

Davidson, D. (1975). *Semantics for Natural Languages,* in D. Davidson and G. Harman (eds.), *The Logic of Grammar*. Encino, CA: Dickenson Publ.

Davidson, D. (2001). *Inquiries into truth and interpretation.* Oxford: Clarendon Press, 2nd edition.

Deliège, C. (1984). *Les Fondements de la Musique Tonale*. Paris: J. C. Lattès.

Deliège, I. (1987). Grouping conditions in listening to music: an approach to

Lerdahl and Jackendoff's grouping preference rules. *Music Perception,* 4(4), 325-360.

Deliège, I. (1996). Cue abstraction as a component of categorization processes in music listening. *Psychology of Music,* 24, 131-156.

Deliège, I. and Sloboda, J. A. (eds.), (1997). *Perception and Cognition of Music*. East Sussex: Psychology Press.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics,* 28(5), 381-389.

Deutsch, D. (1982). The processing of pitch combinations, in D. Deutsch (ed.), *Psychology of Music.* New York: Academic Press, Chapter 9.

Deutsch, D. (1999). *The Psychology of Music*. San Diego: Academic Press.

Deutsch, D. and Feroe, J. (1981). The internal representation of musical pitch sequences in tonal music. *Psychological Review,* 88(6), 503-522.

DeWitt, L. and Samuel. A. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology*, 119(2), 123-144.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *MusicPerception,* 12(1), 1-25.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interaction, in N.L. Wallin, B. Merker and S. Brown (eds.), *The Origins of Music.* Cambridge, MA: MIT Press. pp, 389-407.

Dowling, W.J. (1984). Assimilation and tonal structure: comment on Castellano, Bharucha and Krumhansl. *Journal of Experimental Psychology,* 113(3), 417-420.

Dowling, W. J. (1994). Melodic contour in hearing and remembering melodies, in R. Aiello and J. A. Sloboda (eds.), *Musical Perceptions.* New York: Oxford University Press, 171-190.

Dowling, W. J. (1999). The development of music perception and cognition, in D. Deutsch (ed.), *The Psychology of Music*, Second Edition. San Diego, CA: Academic Press. pp. 603-625.

Dowling, W. J. and Harwood, D. L. (1986). *Music Cognition.* San Diego: Academic Press.

Duchez, M. E. (1979). La representation spatio-verticale du caractere musical grave-aigu et l'elaboration de la notion de hauteur de soin dans la conscience musicale occidentale. *Acta Musicologica,* LI, FASC. I, 34-73.

Edelman, G. (1992). *Bright Air, Brilliant Fire: On the Matter of the Mind.* New York: Basic Books.

Edelman, G. and Tononi, G. (2000). *Consciousness, How Matter Becomes Imagination.* London: Penguin Books.

Epstein, D. (1995). *Shaping Time: Music, the Brain, and Performance.* New York: Schirmer Books.

Fauconnier, G. (1990). Domains and collections. *Cognitive Linguistics,* 1(1), 151-174.

Fauconnier, G. (1997). *Mappings in Thought and Language.* Cambridge: Cambridge University Press.

Feldman, J. and Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language,* 89, 385-392.

Fernandez-Duque, D. and Johnson, M. (1999). Attention metaphors: how metaphors guide the cognitive psychology of attention. *Cognitive Science,* 23(1), 83-116.

Fodor, J. A. and Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 4, 414-420.

Forte, A. and Gilbert, S.E. (1982). *Introduction to Schenkerian Analysis*. New York: W.W. Norton & Company, Inc.

Fraisse, P. (1974). *La Psychologie du Rythme*. Paris: PUF.

Gabrielsson, A. (1993). The complexities of rhythm, in T.J. Tighe and W.J. Dowling (eds.), *Psychology and Music. The understanding of Melody and Rhythm.* Hillsdale: Lawrence Erlbaum Associates. pp. 93-120.

Gabrielsson, A. (2001). Emotions in strong experiences with music, in P. Juslin, and J. A. Sloboda (eds.), *Music and Emotion. Theory and Research.* Oxford: Oxford University Press. pp. 431-452.

Gärdenfors, P. (1988). Semantics, conceptual spaces and the dimensions of music. *Acta Philosophica Fennica,* 43.

Gibbs Jr., R. (1994). *The Poetics of Mind*. Cambridge: Cambridge University Press.

Gibbs Jr., R.W. (1996). Why many concepts are metaphorical. *Cognition,* 61, 309-319.

Gibbs Jr., and O'Brien, J. (1990). Idioms and mental imagery: the metaphorical motivation for idiomatic meaning. *Cognition,* 36, 35-68.

Gibbs Jr., R, Beitel, D. A., Harrington, M. and Sanders, P. (1994). Taking a stand on the meanings of stand: bodily experience as motivation for polysemy. *Journal of Semantics,* 11, 231-251.

Gibbs Jr., R. and Colston, H. (1995). The cognitive psychological reality of image-schemas and their transformations. *Cognitive Linguistics,* 6(4), 347-378.

Gibson, E. and Pearlmutter, N. (2000). Distinguishing serial and parallel parsing. *Jounral of Psycholoinguistic Research,* 29(2), 231-240.

Gjerdingen, R. (2001). An Experimental Music Theory?, in N. Cook and M. Everist (eds.), *Rethinking Music.* Oxford: Oxford University Press. pp.161-170.

Gopnik, A. and Meltzoff, A. N. (1992). Categorization and naming: Basic-level sorting in 18-month-olds and its relation to language. *Child Development*, 63, 1091-1103.

Gregory, A. H. (1978). Perception of clicks in Music. *Perception & Psychophysics*, 24, 171-174.

Guck, M. A. (1991). Two types of metaphoric transfer, in, J.C. Kassler (ed.), *Metaphor: A Musical Dimension.* Sydney: The Currency Press.

Guck, M. A. (1994).Analytical fictions, *Music Theory Spectrum,* 16(2), 217-230.

Hargreaves, D. J. (1986). *The Developmental Psychology of Music*. Cambridge: Cambridge University Press.

Holmes, V. M. and Forster, K. I. (1970). Detection of extraneous signals during sentence recognition. *Perception & Psychophysics*, 7, 297-301.

Horton, T. (1999). *Some Formal Problems with Schenkerian Representations of Tonal Structure*. Unpublished MPhil thesis. Cambridge, UK: University of Cambridge.

Horton, T. (2001). *The Compositionality of Tonal Structures: a Generative Approach to the Notion of Musical Meaning.* Faculty of Music. Cambridge, UK: University of Cambridge.

Horton, T. (2003). *The Formal Structure of Tonal Theory.* Unpublished Doctoral thesis. Faculty of Music, University of Cambridge, UK.

Howell, P., Cross, I. and West, R. (eds.), (1985). *Musical Structure and Cognition.* London: Academic Press.

Howell, P., West, R. and Cross, I. (eds.), (1991). *Representing Musical Structure.* London: Academic Press.

Huron, D. (1991). The avoidance of part-crossing in polyphonic music: perceptual evidence and musical practice. *Music Perception,* 9(1), 93-104.

Huron, D. (2001). Tone and voice: a derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19(1), 1-64.

Imberty, M. (1981). *Les écritures du temps.* Paris: Dunod.

Imberty, M. (1987). De la perception du temps musical à sa signification psychologique: à propos de "la Cathédrale engloutie" de C. Debussy. *Analyse Musicale*, 6, 28-37.

Imberty; M. (1990). A genese des schemes d' orgasation temporelle de la pensee musicale chez l' enfant. *Les Sciences de l' Education,* 3-4, 39-61.

Imberty, M. (1991). Comment l' auditeur organize-t-il; la progression temporelle d'une oeuvre musicale. *Psychologica Belgica*, XXI(2), 181.

Imberty, M. (1992a). De quelques problemes methodologiques dans l'approoche psychologique experimentale de l'interpretation musicale, in R. dalmente and M. Baroni

(eds.), *Secondo Convegno Europeo di Analisi Musicale*. Trento: Universita degli Studi di Trento. pp. 41-49.

Imberty, M. (1992b). Stabilité et Instabilité: comment láuditeur organise-t-il la progression d'une æuvre musicale. In R. Dalmonte e M. Baroni (a cura di) *Secondo Convegno Europeo di Analisi Musicale*. Trento: Universita' degli Studi di Trento. 91-104.

Imberty, M. (1993). Le style musical et le temps: aspects esthésiques et aspects poïétiques. *Analyse musicale*, 32, 14-19.

Imberty, M. (1998). Du vide à l'infini: homologies structurales. *Musicae Scientiae*. Special issue, 91-113.

Jackendoff, R. (1987). *Consciousness and the Computational Mind.* Cambridge, MA: MIT Press.

Jackendoff, R. ([1992]-1999). *Languages of the Mind.* Cambridge, MA: MIT Press.

Jackendoff, R. and Lerdahl, F. (2006). The capacity for music: what is it, and what's special about it? *Cognition,* 100(1), 33-72.

Johnson, K E, and Mervis, C B, (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General,* 126, 248-277.

Johnson, M. (1987). *The Body in the Mind: the Bodily Bases of Meaning, Imagination and Reason*. Chicago: University of Chicago Press.

Johnson, M. (1989). Image-schematic bases of meaning. *RSS,* 9(1-2-3).

Johnson, M. (1993). Conceptual metaphor and embodied structures of meaning: a reply to Kennedy and Vervaeke. *Philosophycal Psychology,* 6, 4.

Johnson, M. (1995). Why metaphor matters to philosophy. *Metaphor and Symbolic Activity,* 10(3), 157-162.

Johnson, M. and Lakoff, G. (2002). Why cognitive linguistics requires embodied realism. *Cognitive Linguistics,* 13(3).

Johnson, M. and Larson, S. (2003). "Something in the way she moves": metaphors of musical motion. *Metaphor and Symbol*, 18(2), 63-84.

Jones, M. R. (1992). Attending to musical events, in M. R. Jones and S. Holleran (eds.), *Cognitive Bases of Musical Communication*. Washington: American Psychological Association, 91-110.

Jones, M. R. (1993). Dynamics of musical patterns: how do melody and rhythm fit together? In T. J. Tighe & W. J. Dowling (Eds.). *Psychology and Music. The Understanding of Melody and Rhythm*. 67-92.

Jones, M. R. and Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*. 96(3), 459-491.

Juslin, P, N. and Sloboda, J. A. (2001). *Music and Emotion. Theory and Research.* Oxford: Oxford University Press.

Kaminska, Z. and Mayer, P. (1993). Transformation, migration and restoration. Shades of illusion in the perception of music. *Contemporary Music Review*, 9(1-2), 151-161.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science.* Cambridge, MA/London: MIT Press.

Kassler, J.C. (1983). Heinrich Schenker's epistemology and philosophy of music: an essay on the relationships between evolutionary theory and music theory, in D. Oldroyd and I.

Langham (eds.), *The Wider Domain of Evolutionary Thought.* Dordrecht, Holland: D. Reidel Publishing Co. pp. 221-260.

Keiler, A. (1978). The empiricist illusion: Narmour's *Beyond Schenkerism. Perspectives of New Music,* Fall-Winter, 161-195.

Kemper, S. (1989). Priming the comprehension of metaphors. *Metaphor and Symbolic Activity,* 4(1), 1-17.

Kielian-Gilbert, M. (2003). Interpreting Schenkerian prolongation. *Music Analysis*, 22(i-ii), 51-104.

Komar, A. (1988). The pedagogy of tonal hierarchy. *In Theory Only,* 10(5), 23-28.

Komar, A. (1992). *Linear-Derived Harmony.* Cincinnati: Ovenbird Press.

Korsyn, K. (2004). The death of musical analysis? The concept of unity revisited. *Music Analysis,* 23(ii-iii), 337-351.

Kramer, J. (2004). The concept of disunity and musical analysis. *Music Analysis,* 23(ii-iii), 361-372.

Kreitzer, A. (1997). Multiple levels of schematization: a study in the conceptualization of space. *Cognitive Linguistics,* 8(4), 291-325.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* 11, 346-374.

Krumhansl, C.L. (1983). Perceptual structures for tonal music. *Music Perception,* 1, 28-62.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

References

Krumhansl, C.L. (1991). Music psychology: tonal structures in perception and memory. *Annual Review of Psychology,* 42, 277-303.

Krumhansl, C. L. and Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review,* 89(4), 334-368.

Krumhansl, C. L. and Shepard, R. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of experimental psychology,* 5(4), 579-594.

Krumhansl, C. L., Bharucha, J. J. and Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Perception and Psychophysics,* 32, 96-108.

Lakoff, G. ([1987]-1990). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind.* Chicago: The University of Chicago Press.

Lakoff, G. (1990). The invariance hypothesis: is abstract reason based on image-schemas? *Cognitive Linguistics,* 1(1), 39-74.

Lakoff, G. (1993). The contemporary theory of metaphor, in A. Ortony (ed.), *Metaphor and Thought.* Second edition. New York: Cambridge University Press.

Lakoff, G. (1994). What is a conceptual system?, in W. Overton and D. Palermo (eds.), *The Nature and Ontogenesis of Meaning.* NJ: Laurence Erlbaum Ass. Chapter 3.

Lakoff, G. and Johnson, M. ([1980] - 2003).*Metaphors we Live By.* Chicago: University of Chicago Press.

Lakoff, G. and Johnson, M. (1999). *Philosophy in the Flesh*. New York: Basic Books.

Lamont, A. & Cross, I. (1994). Children's cognitive representations of musical pitch. *Music Perception*, 12, 27-55.

Lamont, A. and Dibben, N. (2001). Motivic structure and the perception of similarity. *Music Perception,* 18(3), 245-274.

Larson, S. (1997a). The problem of prolongation in tonal music: terminology, perception, and expressive meaning. *Journal of Music Theory*, 41(1), 101-136.

Larson, S. (1997b). Continuations as completions: studying melodic expectation in the creative microdomain 'Seek Well', in M. Leman (ed.), *Music, Gestalt, and Computing. Studies in Cognitive and Systematic Musicology.* Berlin: Springer-Verlag.

Larson, S. (1997-1998). Musical forces and melodic patterns. *Theory and Practice,* 22-23, 55-71.

Larson, S. (2004). Musical forces and melodic expectations: comparing computer models and experimental results. *Music Perception,* 21(4), 457-498.

Lashley, K. S. (1951). The problem of serial order in behaviour, in L. A. Jeffres (ed.), *Cerebral Mechanisms in Behavior*. New York: John Willey & Sons.

Lerdahl, F. (1988). Tonal pitch space. *Music Perception,* 5(3), 315-350.

Lerdahl, F. (1991). Underlying musical schemata, in P. Howell, R. West, and I. Cross (eds.), *Representing Musical Structure.* London: Academic Press Limited, Chapter 8.

Lerdahl, F. (1997). Issues in prolongational theory: a response to Larson. *Journal of Music Theory*, 41(1), 141-155.

Lerdahl, F. (2001). *Tonal Pitch Space*. New York: Oxford University Press.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music.* Cambridge, Massachussets: The MIT Press.

Lewis, D. (1972). General semantics, in D. Davidson and G. Harman (eds.), *Semantics of Natural Language.* Dordrecht: D. Reidel.

Littlefield, R. and Neumeyer, D. (1992). Rewriting Schenker: narrative-history-ideology. *Music Theory Spectrum,* 14(1), 38-65.

Lochhead, J. (1989/90). The metaphor of musical motion: is there an alternative? *Theory and Practice,* 14/15, 83-104.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review,* 95, 492-527.

Logan, G. D. (1992). Attention and preattention in theories of automaticity. *American Journal of Psychology*, 105, 317-339.

Longuet-Higgins, H. C. (1987). *Mental Processes: Studies in Cognitive Science.* Cambridge, MA: MIT Press.

Malloch, S. (2002). Musicality: the art of human gesture, in K Stevens, D. Burham, G. McPherson, E. Schubert and J. Renwick (eds.), *Proceedings of the 7th International Conference of Music Perception and Cognition.* Sydney: University of Western Sydney. 143-146.

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Martínez, I. C. (1997). The transcription of pitches in tonal melodies. A comparative study, in A. Gabrielsson (ed.), *Proceedings of the Third ESCOM Conference.* Uppsala, Sweden: Upssala University, 269-274.

Martínez, I. C. (2000a). La asimetría al juzgar la similitud perceptiva de melodías, in S. Malbrán and F. Shifres (eds.), *Anales de la III Conferencia Iberoamericana de Investigación Musical*. Mar del Plata, Argentina: Conservatorio Luis Gianneo, 85-88.

Martínez, I. C. (2000b). El contraste de rasgos en los juicios de similitud perceptiva entre melodías. Paper presented at the *Encuentro 2000 de Investigación en Arte y Diseño.* Facultad de Bellas Artes, La Plata, Argentina: Universidad Nacional de La Plata.

Martínez, I. C. (2001). Incidence of the context in the perceptual similarity of melodies, in D. Hirts (ed.), *Mikropoliphonie.* http://farben.latrobe.edu.au/mikropol/.

Martínez, I. C. and Shifres, F. (1999a). Utilización de la estructura jerárquica de melodías en juicios de similitud. *Boletín de Investigación Educativo-Musical*, 18, 42-45.

Martínez, I. C. and Shifres, F. (1999b). Music education and the development of structural hearing: a study with children, in M. Barrett; G. Mc Phearson and R. Smith (eds.), *Children and Music: Developmental Perspectives. Proceedings of the Second International Musical Education Research Symposium.* Tasmania, Australia: University of Tasmania, 184-190.

Martínez, I. C. and Shifres, F. (1999c). The rivalry between structure and surface while judging the similarity of melodies. Paper presented at the *1999 SMPC Conference.* Evanston, Illinois, USA: Northwestern University.

Martínez, I. C. and Shifres, F. (2000). Testing models as predictors of the rivalry between structure and surface in the perception of melodies, in C. Woods, G. Luck, R. Brochard, F. Seddon and J. A. Sloboda (eds.), *Proceeding of the Sixth International Conference on Music Perception and Cognition*. Keele, UK: Keele University.

Marvin, E. W. and Laprade, P. (1987). Relating musical contours. Extensions of a theory for contour. *Journal of Music Theory,* 31, 225-267.

Maus, F. E. (1988). Music as drama. *Music Theory Spectrum,* 10, 56-73.

McAdams, S. and Bigand, E. (1994). *Penser les Sons. Psychologie Cognitive de l'Audition.* Paris: Presses Universitaires de France.

McClelland, J.L. and Rumelhart, D.E. (eds.). (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 2. Psychological and Biological Models.* Cambridge, MA: MIT Press.

Medin, D. L. and Ross, B. H. (1997). *Cognitive Psychology.* Orlando: Harcourt Brace College Publishers.

Meltzoff, A. N (1981). Imitation, intermodal co-ordination and representation in early infancy, in: G. Butterworth (ed.), *Infancy and Epistemology.* London: Harverster Press. pp. 85-114.

Mervis, C. and Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology,* 32, 89-115.

Meyer, L. B. (1956). *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Meyer, L. (1973). *Explaining Music. Essays and Explorations.* Chicago: The University of Chicago Press.

Miall, D. and Dissanayake, E. (2003). The poetics of babytalk. *Human Nature,* 14(4), 337-364.

Morgan, R. (2003). The concept of unity and musical analysis. *Music Analysis,* 22(i-ii), 7-50.

Narmour, E. (1977). *Beyond Schenkerism. The Need for Alternatives in Music Analysis.* Chicago: The University of Chicago Press.

Narmour, E. (1983-1984). Some major theoretical problems concerning the concept of hierarchy in the analysis of tonal music. *Music Perception*, 1(2), 129-199.

Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: the Implication-Realization Model.* Chicago: University of Chicago Press.

Narmour, E. (1992). *The Analysis and Cognition of Melodic Complexity.* Chicago: The University of Chicago Press.

Nattiez, J. J. (1992). Existe-t-il des relations entre les diverses méthodes d'analyse?, in R. Dalmonte and M. Baroni (eds.), *Secondo Convegno Europeo di Analisi Musicale.* Trento: Universita degli Studi di Trento. pp. 537-565.

Neisser, U. (1976). *Cognition and Reality*. San Francisco: W. H. Freeman.

Ockelford, A. (2004). On similarity, derivation and the cognition of musical structure. *Psychology of Music,* 32, 23-74.

Ockelford, A. (1991). The role of repetition in perceived musical structures, in Howell, P., West, R. and Cross, I. (eds.), *Representing Musical Structure*. London: Academic Press. pp. 129-160.

O'Donnell, S. (1999). Space, motion, and other musical metaphors, in R. Weigner (ed.), *Perspectives on the Grateful Dead.* Westport, Connecticut: Greenwood Press. pp. 127-135.

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

Pederson, E.; Danziger, E.; Wilking, D.; Levinson, S.; Kita, S., and Senft, G. (1998). Semantic typology and spatial conceptualization. *Language,* 74(3), 557-589.

Peretz, I. and Zatorre, R. (2004). *The Cognitive Neuroscience of Music.* Oxford: Oxford University Press.

Phillips, R.; Wagner, S.; Fells, C., and Lynch, M. (1990). Do infants recognize emotion in facial expressions?: categorical and "metaphorical" evidence. *Infant Behavior and Development,* 13, 71-84.

Pöppel, E. and Wittmann, M. (1999). Time in the mind, in Wilson, R. A. and Keil, F. C. (eds.), MIT Encyclopedia of Cognitive Sciences. Cambridge, MA: MIT Press. pp, 841-843.

Putnam, H. (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press.

Putnam, H. (1990). *Realism with a Human Face*. Cambridge: Harvard University Press.

Quinn, I. (1999). The combinatorial model of pitch contour. *Music Perception,* 16(4), 439-456.

Reybrouck, M. (2001). Musical imagery between sensory processing and ideomotor simulation, in R.I. Godoy and H. Jørgensen (eds.), *Musical Imagery.* The Netherlands: Swets & Zeitlinger. pp. 117-135.

Rosch, E. (1973). Natural cateogories. *Cognitive Psychology,* 4, 328-350.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General,* 104(3), 192-233.

Rosch, E. (1978). Cognitive reference points. *Cognitive Psychology,* 532-547.

Rosch, E. and Mervis, C. B. (1975). Family resemblances. *Cognitive Psychology,* 7, 573-605.

References

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and Boyes-Braen, P. (1976). Basic objects in natural categories. *Cognitive Psychology,* 8, 382-439.

Rosner, B. (1984). A Generative Theory of Tonal Music by Fred Lerdahl and Ray Jackendoff (review). *Music Perception,* 2, 289-290.

Rosner, B. and Meyer, L. (1986). The perceptual roles of melodic process, contour and form. *Music Perception,* 4, 1-40.

Rothfarb, L. (2002). Energetics, in T. Christensen (ed.), *Cambridge History of Western Music Theory.* Cambridge: Cambridge University Press. pp. 927-955.

Rothgeb, J. (1983). Thematic content: a Schenkerian view, in D. Beach (ed.), *Aspects of Schenkerian Theory.* New Haven: Yale University Press, 39-60.

Rumelhart, D. E. and Ortony, A. (1977). The representation of knowledge in memory, in R.C. Anderson, R.J. Spiro, & W.E. Montague (eds.), *Schooling and the Acquisition of Knowledge.* Hillsdale, NJ: Erlbaum.

Saffran, J.; Newport, E., and Aslin, R. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language,* 35, 606-621.

Salzer, F. ([1962]-1982). *Structural Hearing. Tonal Coherence in Music.* New York: Dover Publications, Inc.

Salzer, F. and Schachter, C. [1969] - 1989). *Counterpoint in Composition.* New York: Columbia University Press.

Saslow, J. (1996). Forces, container, and paths: the role of body-derived image-schemas in the conceptualization of music. *Journal of Music Theory,* 40(2), 217-243.

Saslaw, J. (1997-1998). Life forces: conceptual structures in Schenker's *Free Composition* and Schoenberg's *The Musical Idea*. *Theory and Practice*, 22-23, 17-33.

Schachter, C. (1976). Rhythm and linear analysis: a preliminary study. *Music Forum,* 4, 281-334.

Schachter, C. (1980). Rhythm and linear analysis: durational reduction. *Music Forum,* 5, 197-232.

Schachter, C. 1981. A commentary on Schenker's *Free Composition. Journal of Music Theory,* 25(1), 115-142.

Schachter, C. (1990). Either/Or, in H. Siegel (ed.), *Schenker Studies*. Cambridge: Cambridge University Press, 165-179.

Schachter, C. (1998). *Unfoldings*: *Essays in Schenkerian Theory and Analysis*. New York: Oxford University Press.

Schachter, C. (1999). Structure as foreground: "das drama des Ursatzes", in C. Schachter and H. Siegel (eds.). *Schenker Studies 2.* New York: Cambridge University Press, 298-314.

Schenker, H. ([1906]-1990). *Tratado de Armonía* [*Harmonielehre,* trans: R. Barce]. Madrid: Real Musical.

Schenker, H. ([1910-1922]-1987). *Counterpoint*. [Kontrapunkt, trans. J. Rothgeb and J. Thym.]. New York: Schirmer Books.

Schenker, H. ([1925]-1994). *The Masterwork in Music: A Yearbook*; Vol. I. [W. Drabkin (ed.), trans: Ian Bent, William Drabkin, Richard Kramer, John Rothgeb, and Hedi Siegel]. Cambridge: Cambridge University Press.

Schenker, H. ([1926] - 1996). *The Masterwork in Music: A Yearbook*; Vol. II. [W. Drabkin (ed.), trans: Ian Bent, William Drabkin, John Rothgeb, and Hedi Siegel]. Cambridge: Cambridge University Press.

Schenker, H. ([1930] - 1997). *The Masterwork in Music: A Yearbook*; Vol. III. [W. Drabkin (ed.), trans: Ian Bent, Alfred Clayton, Derrick Puffett]. Cambridge: Cambridge University Press.

Schenker ([1933]-1969). *Five Graphic Music Analyses.* New York: Dover Publications, Inc.

Schenker, H. ([1935] - 1979). *Free Composition.* [*Der freie Satz*, trans. E. Oster]. New York: Schirmer Books.

Schenker, H. ([1921 – 1924] 2004 – 2005). *Der Tonwille: Pamphlets' Quarterly Publication in Witness of the Immutable Laws of Music*. 2 vols. [W. Drabkin, ed. and contr. trans.].New York: Oxford University Press.

Schmuckler, M.A. (1999). Testing models of melodic contour similarity. *Music Perception,* 16(3), 295-326.

Schubert, E. (2001). Continuous measurement of self-report emotional response to music, in P. N. Juslin and J. A. Sloboda (eds.), *Music and Emotion. Theory and Research.* Oxford: Oxford University Press, Chapter 17, pp. 393-414.

Scruton, R. (1983). *The Aesthetic Understanding: Essays in the Philosophy of Art and Culture*. London: Methuen.

Selfridge-Field, E. (1998). Conceptual and representational issues in melodic comparison, in W.B. Hewlett and E. Selfridge-Field (eds.), *Computing in Musicology,* 11, 3-64.

References

Serafine, M. L. (1988). *Music as Cognition. The Development of Thought in Sound.* New York: Columbia University Press.

Serafine, M. L., Glassman, N. and Overbeeke, C. (1989). The cognitive reality of hierarchical structure in music. *Music Perception,* 6(4), 397-430.

Shepard, R. N. (1982). Structural representations of music pitch, in D. Deutsch (ed.), *The Psychology of Music.* New York: Academic Press. pp. 344-390.

Shifres, F. and Martínez, I. C. (1999). Control experimental de la estructura jerárquica y la superficie musical. *Boletín de Investigación Educativo-Musical*, 17, 42-46.

Shifres, F. and Martínez, I. C. (2000a). Incidencia del contorno melódico en la representación reduccional en músicos y no músicos, in J. Tafuri (ed.), *La ricerca per la didattica musicale. Quaderni della SIEM*, 16, 294-304.

Shifres, F. and Martínez, I. C. (2000b). The role of performance in the cognitive reality of the hierarchic structure, in C. Woods, G. Luck, R. Brochard, F. Seddon and J. A. Sloboda (eds.), *Proceeding of the Sixth International Conference on Music Perception and Cognition.* Keele, UK: University of Keele.

Shifres, F. and Martínez, I. C. (2000c). Examinando el rol de la ejecución en la cognición de la estructura musical jerárquica. Un estudio basado en tareas de estimación, in S. Malbrán and F. Shifres (eds), *Anales de la III Conferencia Iberoamericana de Investigación Musical.* Mar del Plata, Argentina: Conservatorio Luis Gianneo, 118-122.

Shifres, F. and Martínez, I. C. (2002). Regulación de los procesos de atención a través de la estructura suybyacente de la música tonal. Evidencia ontogénica, in S. Malbrán (ed.), *Anales de la 4ta Conferencia Iberoamericana de Investigación Musical.* Universidad Nacional de San Juan. San Juan: Facultad de Filosofía y Humanidades. pp. 87-100.

Simon, H. A. & Sumner, R. K. (1968). Pattern in Music, in B. Kleinmuntz (ed), *Formal representation of human judgment.* New York: Wiley. Reprinted in S. Schwanauer and D. Levitt (eds) (1993). *Machine Models of Music,* MIT Press: Cambridge, MA, pp. 83-112.

Sloboda, J. A. ([1985]-1996). *The Musical Mind. The Cognitive Psychology of Music.* Oxford: Clarendon Press.

Sloboda, J.A. & Gregory, A. H. (1980). The psychological reality of musical segments. *Canadian Journal of Psychology*, 34, 274-280.

Sloboda, J. A. and Davidson, J. W. (1996). The young performing musician, in I. Deliege and J. A. Sloboda (eds.), *Musical Beginnings: Origins and Development of Musical Competence.* Oxford: Oxford University Press, 171-190.

Sloboda, J. A. and Parker, D. H. H. (1985). Immediate recall of melodies, in P. Howell, I. Cross, and R. West (eds.), *Musical Structure and Cognition.* London: Academic Press.

Snyder, B. (2000). *Music and Memory*. Cambridge, Massachussets: The MIT Press.

Spitzer, M. (2004). *Metaphor and Musical Thought.* Chicago: The University of Chicago Press.

Steinbeis, N. & Koelsch, S. (2007). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex,* doi:10.1093/cercor/bhm149.

Stern, D. (1985). *The Interpersonal World of the Infant. A View form Psychoanalysis and Developmental Psychology.* New York: Basic Books.

Stoffer, T. (1985). Representation of phrase structure in the perception of music. *Music Perception*, 3(2), 191-220.

Temperley, D. (2001). *The Cognition of Basic Musical Structures.* Cambridge, Massachussets: The MIT Press.

Thomas, N. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science,* 23(2), 207-245.

Todd, N. P. (1985). A model of expressive timing in tonal music. *Music Perception*, 3(1), 33-58.

Trevarthen, C. (1999/2000). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae*, special issue, 155-213.

Turner, M. (1990). Aspects of the invariance hypothesis. *Cognitive Linguistics,* 1(2), 247-255.

Turner, M. (1993). An image-schematic constraint on metaphor, in R. Geiger and B. Rudzka-Ostyn (eds.), *Conceptualizations and Mental processing in Language.* Berlin: Mouton de Gruyter.

Tversky, A. (1977). Features of similarity. *Psychological Review,* 84(4), 327-352.

Tversky, B. and Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General,* 113(2), 169-193.

Varela, F., Thompson, E., and Rosch, E. (1993). *The Embodied Mind. Cognitive Science and Human Experience.* Cambridge, MA: MIT Press.

Wagner, S.; Winner, E.; Cicchetti, D. and Gardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development,* 52, 728-731.

Zbikowsky, L. (1997). Conceptual models and cross-domain mapping: new perspectives on theories of music and hierarchy. *Journal of Music Theory,* 41(2), 193-225.

Zbikowsky, L. (1997-1998). Des Herzraums abschied: Mark Johnson's theory of embodied knowledge and music theory, *Theory and Practice,* 22-23, 1-14.

Zbikowsky, L. (1998). Metaphor and music theory: reflections from cognitive science. *Music Theory Online,* 4(1).

Zbikowsky, L. (2002). *Conceptualizing Music.* New York: Oxford University Press.