A Fundamental Explanation of Musical Meaning in Terms of Mental States

by
Barry Ross

Thesis presented in partial fulfilment of the requirements for the degree Master of Music at the University of Stellenbosch

Supervisor: Prof. Winfried Lüdemann
Co-supervisor: Dr. Johannes P. Smit
Faculty of Arts and Social Sciences
Department of Music

December 2010

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2010

Copyright © 2010 University of Stellenbosch

All rights reserved

ABSTRACT

This study concerns the widespread phenomenon that music is perceived as meaningful to the listener in some sense. The study adopts a style of conceptual clarification and investigation that is current in the analytic philosophy of language, and is further informed by recent research into the fundamental biology of human musicality, which suggests that musicality and language are neurologically related.

The problem of musical meaning is approached as a problem of communicative behaviour, and is hence conceptually related to the concept of meaningfulness in the various modalities of linguistic communication. 'Communication' is defined in terms of the intended consequences of communicative acts – that is, a communicative act is an attempt on the behalf of the utterer to cause some sort of change in the listener's mental states. From this premise, meaning in both musical and linguistic acts is defined in terms the mental states elicited in the mind of the listener. Two classes of mental state are identified: *cognitive states*, which are propositional in nature; and *affective states*, which are essentially non-propositional. It is proposed that meaning in both music and language (as well as in other communicative acts) can be explained in terms of the elicitation of these classes of mental states in the minds of competent listeners, and that in any linguistic or musical act, a competent listener will entertain a composite of these mental states that will be perceived as meaning.

The mechanisms responsible for the elicitation of these states are discussed, and it is concluded that the causal powers of the communicative act, as it is represented in the mind, are responsible for the elicitation of these mental states. Directly causal means are responsible for affective states: there is a relationship of direct causation between relevant features of the communicative act, as represented in the mind, and affective states. Affective states are non-propositional, in that they cannot be subjected to deductive or propositional operations in the mind. By virtue of their being non-propositional, such states are also considered to be beyond verbal explication ('ineffable'). Cognitive states, on the other hand, are propositional in nature. The mechanisms by which they are realised are complex in terms of propositional computation: the relevant propositional features of the communicative act, as represented in the mind of the listener, undergo manipulation by mental processes (for instance, the computational system for linguistic syntax). Cognitive states are expressible in propositional terms, and are hence expressible in language.

Whereas linguistic communication is efficacious for the elicitation of cognitive states, musical utterances tend to elicit affective states to a far greater degree. Furthermore, whereas the syntax of language aids communication in the facilitation of semantics, the syntactic dimension of music is principally a means of implementing affective states in the listener. Therefore, any explanation of musical meaning must take the syntactical dimension of music into account. It is also argued that there are features of performance common to both language (in its spoken modality) and musical utterances that serve to elicit affective states.

OPSOMMING

Hierdie studie ondersoek die verskynsel dat musiek deur die meeste luisteraars as betekenisvol ervaar word. ’n Styl van konseptuele verduideliking en ondersoek word gebruik wat eie is aan die analitiese filosofie van taal. Terselfdertyd word die jongste navorsing op die gebied van die fundamentele biologie van menslike musikaliteit in aanmerking geneem, wat suggereer dat taal en musikale vermoë neurologies met mekaar verwant is.

Die probleem van betekenis in musiek word as ’n probleem van kommunikatiewe gedrag benader, en is dus konseptueel verbind aan die konsep van betekenisvolheid in die verskeie modaliteite van kommunikasie deur middel van taal. ‘Kommunikasie’ word in terme van die geïntendeerde uitkomste van kommunikatiewe aksies/dade gedefinieer. Met ander woorde, ’n kommunikatiewe aksie/daad is ’n poging deur die spreker om uiteindelik ’n verandering in die geestesgesteldheid (‘mental state’) van die luisteraar teweeg te bring. Op hierdie basis word twee tipes geestesgesteldheid onderskei: ’n kognitiewe gesteldheid, wat proposisioneel van aard is, en ’n affektiewe gesteldheid, wat nie-proposisioneel is. Daar word voorgestel dat betekenis in beide musiek en taal, soos ook in ander vorme van kommunikasie, verduidelik kan word as die belewenis van sodanige geestesgesteldhede aan die kant van die bedrewe luisteraar. Dit impliseer dat die betekenis van enige uiting in taal of musiek as ’n bepaalde kombinasie van hierdie twee geestesgesteldhede deur die bedrewe luisteraar ervaar word.

Die meganismes wat hierdie geestesgesteldhede ontlok word bespreek, en die gevolgtrekking word gemaak dat dit die kousale mag van die kommunikatiewe daad is, soos dit in die bewussyn (‘mind’) neerslag vind, wat hierdie twee tipes geestesgesteldheid ontlok. Daar word beweer dat ’n proses van direkte kousaliteit verantwoordelik is vir ’n affektiewe gesteldheid: daar is ’n oorsaaklike verhouding tussen die onderskeie kenmerke van die kommunikatiewe daad, soos dit in die bewussyn voorgestel word, en die uiteindelike affektiewe geestesgesteldheid. ’n Affektiewe geestesgesteldheid is nie-proposisioneel omdat dit nie in terme van deduktiewe of proposisionele prosesse in die bewussyn verstaan kan word nie. Omdat dit nie-proposisioneel is word die kenmerke van hierdie affektiewe geestesgesteldheid as onsegbaar (‘ineffable’) deur die luisteraar beleef. Daarteenoor is ’n kognitiewe geestesgesteldheid proposisioneel van aard. Die meganismes wat veroorsaak dat hierdie geestesgesteldheid gerealiseer word is kompleks: die onderskeie kenmerke van die kommunikatiewe daad, soos dit in die bewussyn van die luisteraar voorgestel word, ondergaan manipulasie deur denkprosesse wat proposisioneel van aard is (bv., die denkproses wat die sintaktiese dimensie van taal moet verwerk). ’n Kognitiewe geestesgesteldheid kan in proposisionele terme weergegee en gevolglik in taal verwoord word.

Terwyl kommunikasie deur middel van taal effektief is om ’n kognitiewe geestesgesteldheid te ontlok, is musikale uitdrukking veel eerder geskik om ’n affektiewe geestesgesteldheid te ontlok. Verder, terwyl die sintaksis van taal bydra tot verwesenliking van semantiese betekenis, dra die sintaktiese dimensie van musiek eerder daartoe by om ’n affektiewe geestesgesteldheid by die luisteraar te vestig. Dus moet elke verduideliking van musikale betekenis die sintaktiese dimensie van musiek in aanmerking neem. Verder word beweer dat daar algemene kenmerke in sowel taal (in die gesproke modaliteit) as musiek voorkom wat spesifiek ’n affektiewe geestesgesteldheid tot stand bring.

ACKNOWLEDGEMENTS

I am indebted to my supervisors, Prof. W. A. Lüdemann and Dr. J. P. Smit, for their invaluable input and support during this study. Without the knowledge and advice that they imparted, the completion of this project would surely have been impossible.

CONTENTS

DECLARATION **1**

ABSTRACTS **2**

ACKNOWELDGEMENTS **4**

CONTENTS **5**

CHAPTER 1: INTRODUCTION TO THE STUDY **8**

1.1 Introduction 8
1.2 The meanings of ‘meaning’: arbitrary signs, indexes and icons; ‘meaning’ as significance 12
1.3 Fundamental assumptions: the ‘matter of the mind’ 15
1.4 Linguistic modality and linguistic computation 20

CHAPTER 2: INTRODUCTION TO COGNITIVE AND AFFECTIVE MENTAL STATES **24**

2.1 Introduction: A model of communication 24
2.2 Pragmatic context in linguistic communication 29
2.3 Music and language as communicative media 33
2.4 Mental states as a precondition for musical meaning 37
2.5 Causal mechanisms surrounding the elicitation of appropriate mental states 40
2.6 The argument so far 46

CHAPTER 3: STRUCTURE, SYNTAX, SEMANTICS, AND CONTEXTUAL MEANING IN MUSIC AND LANGUAGE **48**

3.1 Introduction 48
3.2 Structure in music and language 49
3.3 Semantics in music and language 53
3.4 Non-communicative contextual 'meaning' 55
3.5 Some more comments about syntax 60

CHAPTER 4: CAUSAL MECHANISMS AND AFFECTIVE STATES IN MUSIC AND LANGUAGE 64

4.1 Introduction 64
4.2 Affective states and language 66
4.3 Affective states and music 70
4.4 More detail about parts of musical utterances resulting in affective states 74
4.5 A brief discussion of ineffability and musical statements 81

CHAPTER 5: MUSICAL SYNTAX AND THE ELICITATION OF MENTAL STATES 86

5.1 Introduction: The question of musical universals 86
5.2 Form and syntax 91
5.2.1 Formal and syntactical analyses of phrases 91
5.2.2 Formal and syntactical analyses of structures larger than the phrase 98
5.3 The robustness of the notion of 'syntax' in musical phrases 99
5.4 Syntax and the formation of cognitive and affective states 105
5.5 Mechanisms of causality 108
5.5.1 Directly causal mechanisms and cognitive-causal links in language 108
5.5.2 The elicitation of cognitive states from musical utterances 109

CHAPTER 6: SUGGESTIONS REGARDING THE FUNDAMENTAL NATURE OF MUSICAL SYNTAX 116

6.1 Introduction 116
6.2 Chomsky's fundamental assumptions and their applicability to the GTTM and cognitive theories of music perception 121
6.3 Groups as formatives in musical syntax: a disanalogy between music and language 129
6.4 Grouping 132
6.5 The question of narrative or coherence rules above the level of the phrase 142
6.6 Overview: Basis of a theory of syntax 145

CHAPTER 7: CONCLUSION 148

7.1 Introduction 148
7.2 Key differences in approach 150
7.2.1 The term 'semantics' 150
7.2.2 Ontological questions 151
7.2.3 Avoidance of arguments claiming that meaning is the perception of some specific feature of musical experience 152
7.2.4 Universality and ethnocentricity 153
7.3 Is 'meaning' still missing from the computational machine? 155
7.4 Conclusion 160

REFERENCE LIST 162

CHAPTER 1
INTRODUCTION TO THE STUDY

1.1 Introduction

The principal inspiration for this study is the recent resurgence in interest regarding the origins of human music. Over the past decade, this topic has drawn attention from numerous researchers in varied fields, including musicology, ethnomusicology, cognitive musicology, palaeoanthropology, archaeology, the cognitive neurosciences and brain studies, linguistics, infant studies, and animal ecology. The interest in the topic is largely fuelled by contemporary work in cognitive neuroscience and brain studies, which has challenged the assumption that human musical behaviour is the product of social forces alone. Research around musical issues in cognitive neuroscience appears to suggest that there is a large biological component to the story of human music, and current empirical and theoretical studies have only served to further reinforce this idea. The interest in evolutionary questions follows logically: if musical behaviour is to be regarded as having a biological manifestation, one would expect this biological manifestation to have an evolutionary history.

Enough interest in the topic has been generated that any attempt to list academic contributions to the debate would run into many pages. Approaches to the topic abound, and all approaches raise questions about human musicality. What precisely is this thing we call 'music'? In which ways is it biological? Is it related to other capacities, such as language, and if so, in what ways? Can aspects of human musicality be regarded as evolutionary adaptations that aided survival in some manner, or are they spin-offs of other capacities? Are there physical proxies in the archaeological record that can serve as useful markers for the course of the evolution of human musicality? If human musicality can be said to have an evolutionary history, was it the case that music existed in some more basic state, a sort of 'proto-music'? And what could such a proto-music have been used for?

Academic circumstance is not without its cases of irony, for it was linguist Steven Pinker's assertion that music constitutes nothing more than "auditory cheesecake", pleasurable but unessential to survival, that served as an early spark to the debate about the origins of music (Pinker, 1997:529-539). This is not to say, however, that Pinker's view of music being mostly

parasitic with regard to neural hardware is entirely incorrect. There currently exist several theories within evolutionary musicology that see music as being somehow related to the human capacity for language, and hence parasitic on the language faculty's neural hardware to a certain degree. Music-language comparisons in general have long been a topic of interest and vigorous debate, with different sides arguing about why it is that music should or should not be regarded as some sort of language. However, even a superficial consideration of the analogy is bound to lead one to realise that while music may appear to be language-like in many respects, it is clearly different in others. To this day, diverging views exist on music's status as a language or as having language-like qualities. For example, in a book entitled *Musical Languages* (1997), musicologist Joseph Swain argues that music is akin to a sort of language, while the philosopher Roger Scruton, in a book published in the same year entitled *The Aesthetics of Music*, dedicates entire chapters to argue that it is not.

These two authors were not, however, interested in the origins of musical behaviour. Neither was their interest principally in the data surrounding the biological grounding of human musicality. There are numerous examples of studies predicated on the idea that, at the very least, music and language are *conceptually* analogous. What the data surrounding the biology of musicality suggests, however, is that the human capacities for music and language are to some extent biologically linked as well. This study will be an attempt at a synthesis of these views: that thinking about an analogy between music and language is, for better or worse, a profitable conceptual enterprise; and that the capacities for music and language are biologically linked in some way, and thus to some extent share physical manifestation. The crossroads at which these two approaches will meet in this study is the question of precisely how it comes to be that music is perceived by listeners as being somehow *meaningful*.

The question of musical meaning may seem far removed from evolutionary musicology. However, with data existing suggesting that music and language are linked in some respect, questions about how music and language differ is also of interest to researchers wishing to account for the evolutionary history of either behaviour. Within the corpus of research on and relevant to music origins, the question of musical meaning has found serious consideration into two sets of work: a series of articles published by musicologist Ian Cross (e.g. Cross, 2005), and a chapter in a major work on the neurobiology and evolutionary history of music by Aniruddh Patel (2008a). Both authors agree that the common human apprehension of music as meaningful requires explanation. Cross's approach is interesting in that it is a novel

conceptual model of meaning being used to account for the functionality of musical behaviour in both the present time and possibly the distant past; Patel's approach is interesting in that it takes an empirical approach to the question of musical meaning, testing the lengths to which the music-language analogy can be stretched with scientific methodology.

It is precisely here, with the question of meaning, that another academic discipline may come to contribute something to the debate about the biological heritage of musicality. Absent from the list of disciplines above was philosophy. In particular, it is within the analytic philosophical tradition, and specifically the analytic philosophy of language, that one would expect to find interest in the perplexities of meaning in human communication. Since the turn of the last century, the predominant form of philosophy practiced in the English-speaking world has been of the analytic variety. This field owes its ascendancy to a general disillusionment with idealist philosophies, such as that of Hegel. Coupled with an interest in logic and the foundations of mathematics, philosophers such as Bertrand Russell and Gottlieb Frege became interested in technical matters surrounding language use, especially in logical terms. Anglophone philosophy has remained interested in questions about the logic of language, and has also reached out into questions about the nature of the mind in relation to the brain, with consciousness and the mechanisms of thought becoming principal concerns in this area. Combined with a more recent return to traditional problems of philosophy, such as ethics, politics, and the nature of reality, the majority of philosophy departments in the English-speaking world can be classified as analytic.

The philosophy of the European-influenced non-English speaking lands (including predominantly Afrikaans South African university departments) has tended toward what is termed continental philosophy. Here, philosophical topics of wider cultural relevance are usually discussed, such as justice, ethics, metaphysics and religion, history, politics, and so on. This field was also subject to the immense interest in language around the beginning of the 20th century. However, the focus of interest regarding language in continental philosophy is in the interpretation of texts, the use of language in situations of power, language's role in the constitution of social reality, and so on. In analytic philosophy, the interest in language is narrower, with the relation between cognition and language, the nature of meaning and its manifestation in cognition, and the nature of truth in language. However, what most notably distinguishes analytic philosophy from its counterpart is a matter of style. The problems of

analytic philosophy are technical, and attention is paid to detail rather than the generalising and sometimes sweeping philosophical systems found in continental philosophy. At the beginning of the twentieth century, many analytic philosophers believed that a great many philosophical problems arise from confused language use, and hence faulty logic. Whereas continental philosophy sees itself as fundamental to human knowledge in general, the first analytic philosophers often considered their task as the "logical clarification of thoughts",[1] and philosophy as subservient (as opposed to fundamental) to the natural sciences. Today, the diversity of topics covered in analytic philosophy is far wider, with language, logic, science and the philosophy of mind being represented in equal proportions to politics, metaphysics, ethics, aesthetics, and other traditional philosophical topics. However, the technical and rigorous style of inquiry, along with the subservient attitude toward science, has remained.

What can analytic philosophy offer questions concerning the origins of music? As one can imagine, the awareness of conceptual issues that is present in much analytic philosophy is of obvious use. However, our own concern will be with the nature of meaning in language and music, approached from a cognitive perspective. It is often claimed that music is somehow meaningful. As has already been noted, the question of the role of musical meaning has found its way into literature about the origins of music. What is this musical meaning, and how is it distinguishable from the meaning that is characteristic of language? If language arose in our species as a means of communication or conveying meaning, how does music fit into the picture? What sorts of mechanisms of cognition are responsible for the perception of musical meaning? How is it that a musical stimulus results in this mental phenomenon? And if music and language share an evolutionary heritage, what role could meaning potentially have played?

Before our understanding of the potential evolutionary significance of musical meaning can be properly ascertained, we must identify what this thing called musical meaning is. Furthermore, we must do so in a way that will allow us to, at least in principle, begin to approach an empirically relevant theory of what it is that happens when people describe a piece of music as meaningful. To do so, this study will consider what sorts of mental states are elicited in the minds of listeners of musical pieces that are perceived as meaningful ('meaning' being understood only in the communicative sense). Also considered will be the

[1] Wittgenstein, *Tractatus Logico-Philosophicus*, §4.112 (1961:29).

mechanisms by which these meaning-relevant mental states are achieved. In doing so, we will pursue a loose analogy between music and language. This analogy will be cast in terms of the sorts of communication-relevant mental states that are elicited by both behaviours, and the mechanisms by which either behaviour elicits those mental states. It will be suggested that a productive manner of thinking about linguistic and musical meaning in the mind is by thinking of meaning in terms of mental states. In other words, these mental states are constitutive of the perception of meaning in the mind. Meaning in both language and music, as well as other sorts of communicative and interactive activities, is the result of mental mechanisms that constitute certain types of mental states. These states determine what we perceive as 'meaning', whether in language, music, writing, or any other sort of communicative activity. Of course, it is not being suggested that the mental states resulting from language and music are identical; rather, they are of the same class of mental state, with the similar meaning-relevant characteristics.

1.2 The meanings of 'meaning': arbitrary signs, indexes, and icons; 'meaning' as significance

It should already be clear from the discussion above that in this study (and in analytic philosophy) we deal with the notion of 'meaning' in a specified sense. This sense can be broadly classified as communicative in nature. When we speak of the meaning of a sentence or statement in language, we are talking about what is communicated. The model of meaning that is going to be discussed in this study is of a communicative variety, with substantial modifications to our understanding of communication. Therefore, our focus is on meaning as a property of a communicative act: the meaning of words and sentences in the various modalities of language, such as speaking and writing. In a similar manner, other communicative or symbolic systems can have the property of meaning in this sense. Traffic lights, Morse code, and gestures have this sort of communicative meaning. Common to all of these cases of meaning is the fact that the intention to communicate is present.[2] Also evident is the fact that these examples are all cases of symbols, where a symbol (or 'sign') is chosen to stand for (or refer to) something other than itself, to which it is not usually causally

[2] In exceptional circumstances, that intention is not there: for example, a teenager's private diary, which is not intended to be read by anyone at all. These cases are rare, and to counter this line of argument, the notion of communicative meaning is robust enough for present purposes if we simply accept that the potential for intention to communicate is present.

connected.[3] This is the property of arbitrariness: there is nothing intrinsic to the actual symbols for comprising the word 'cat' that is directly linked to the furry four-legged feline that purrs. The relationship between the letters C-A-T (or the sound that you make when you utter the word) is one of arbitrary convention, which has been dictated by a community of symbol-users who have agreed that the word and sound refers to furry four-legged felines. The characteristic of arbitrariness with regard to musical meaning is due to undergo considerable scrutiny in the pages that follow.

Another sort of 'meaning' can be thought of as circumstantial, or consequential. Consider a scenario where a hiker is on a trail, looks to the horizon, and sees a plume of smoke emanating from a far-off clump of dry trees. He may come to the conclusion that there is a fire. The hiker may actually utter the words, "That smoke *means* that there is a fire". This sort of language use is common in similar situations: the heavy clouds mean rain; war means suffering; a broken kettle means no coffee. (A semiotician may be more inclined to say that smoke *signifies* fire.) Clear from these examples is that neither the fire, nor the rain, intends to communicate with you by means of smoke or dark clouds. Rather, these are things that are causally connected with their 'meaning', and that 'meaning' is something that you, as an observer, infer from your environment. Information is gleaned from the environment, and the consequences of what is observed are inferred. Such 'signs' are often referred to as *indexes*.[4] There is, however, not necessarily an intention to communicate. The fire does not try to communicate its presence to you by means of smoke. Of course, one can use indexes as arbitrary symbols in themselves: consider making a fire to send smoke signals. The smoke itself is an index of fire, but it is used as an arbitrary signal with a conventional communicative meaning. In this study, I will essentially argue that something similar to indexes being used to communicate is a part of musical meaning. Of course, it may be the case that we do not perceive the causal connection between the index and what it means, but rather simply associate the sign with something due to past experience.[5]

[3] The following discussion of arbitrary signs, indexes, and icons, is mirrored in concise terms in Cross & Tolbert (2009), which is in turn based on work by the philosopher C. S. Peirce. Other similar (semiotic) approaches can be found in Nattiez (1990; 2004) and Tarasti (1995).

[4] Indexes share a similarity but should not be confused with 'indexicals', which are words in language whose reference depends on the situation of the person who utters them, such as 'I', 'me', 'here', 'now', and so on. Indexicals are a particular topic of interest for analytic philosophers of language.

[5] This sort of meaning by association has been considered with regard to music by Burkholder (2006). He simply calls it "associative musical meaning". At times, his model seems to slip between communicative and non-communicative (significance) senses of 'meaning', a distinction that is discussed in the main text.

Another variety of meaning is direct resemblance to the referent. With these cases, there is some resemblance to the referent reflected in the sign itself. Consider the common Western sign used to mark a bathroom as reserved for ladies: a stick figure with a triangular shape superimposed (🚺). This anthropomorphic shape represents the female gender by virtue of us associating the shape of the triangle with the outline of a skirt, a garment traditionally worn by women. While a community could agree to use this symbol differently, say for men's bathrooms, in a culture where women often wear skirts, it would be difficult to get rid of the symbol's association with the female sex. We are inclined, when presented with this symbol, to say that it means 'woman' by virtue of the fact that the figure appears to be wearing something analogous to a skirt (or 'women's bathroom' due to prior experience of the symbol). A more striking example is that of paintings of people: these are in effect objects that 'refer' to real people. What is important to note is that with these sorts of symbols the connection between sign and referent is not arbitrary, but rather depends on a similarity with regard to vital characteristics. Indeed, it is from depictions of religious figures that this class of sign receives its name: *icons*. Hence, we may be prompted to say that some symbols are iconic in nature, due to the fact that they non-arbitrarily represent their referent.

The three categories of symbols mentioned above – arbitrary symbols, indexes, and icons – are of particular interest to the field of musical semiotics. This is not indicative of the approach that this study will take. What is noteworthy is that these signs all have communicative potential – they can be used to communicate with others. The 'meaning' we attribute to such signs and symbols is of a communicative nature. However, the word 'meaning' is occasionally used in non-communicative senses as well. In order to distinguish our communicative approach to meaning from other senses of the word (that is, our understanding of meaning as a property of communication), we should briefly investigate the other principal use of the word.

The non-communicative sense of the word 'meaning' can be thought of as having a connotation similar to 'significance', and is not the same thing as communicative meaning. Instances of this concept's usage are cases of the 'meaning (or significance) of something *for* someone'.[6] Take the following question: What is the meaning of the Munich Agreement? We

[6] Readers of Cross and Tolbert (2009) may notice the use of 'significance' in a slightly different sense in the opening paragraphs, viz. something has 'significance' if it points to or indicates the existence of something else.

could have a look at the Agreement, and figure out the meaning of the words as instances of communicative meaning (arbitrary symbols, according to our discussion above). However, there is another way to interpret this question: namely, as a question about the *significance* of the Munich Agreement. More accurately, meaning as significance should be thought of as 'meaning for *x*', or 'significance for *x*'. Such meaning is always a meaning *for* someone. Therefore our question about non-communicative meaning could be rephrased as: What is the *significance* of the Munich Agreement *for x*? We may come up with a variety of answers. For Hitler, it meant success; for historians, it meant a failed attempt at appeasement; for contemporary Czechoslovakians, it meant disaster. For British comedy troupe Monty Python, its significance for history was satirised as "Britain's great pre-war joke". For young, modern Europeans, the Munich Agreement may mean little more than an unfortunate fact of history, regardless of whether they are aware of its role in shaping the modern Western world. Questions of meaning as significance are tied closely to questions about value. But none of these 'meanings' or significances are related to the communicative notion of 'meaning', which is a property of communicative acts. The value we place in a piece of literature is not the same as the communicative meaning that is a property of the sentences which comprise it. And while pieces of music may have great 'meaning' in our lives, this significance should not be confused with communicative meaning. In this study, our concern is with meaning as a property of communicative acts. The non-communicative sense of the term 'meaning' (as significance) is *not* the object of investigation in this study. For the remainder of this study, the term 'meaning' will refer to meaning in its communicative sense, and not meaning as 'significance for *x*'.

1.3 Fundamental assumptions: the 'matter of the mind'

One would expect that a theory of musical meaning that has been inspired by interests in the biological constitution of human musicality would be closely allied to a naturalistic conception of the mind. In other words, a naturally congruent view of the mind in evolutionary investigations is that the 'matter of the mind' is the human brain.[7] This is indeed the case in this study, and one of its motivations. In analytic philosophy, naturalism about

---

Cross and Tolbert do not persist with this terminology, and in this study, I have chosen to use the term in the sense specified in the main text.

[7] The term 'matter of the mind' is borrowed from Lewis-Williams (2002).

mental phenomena is common (Blackmore, 2006). This includes accounts of linguistic phenomena – that of Devitt and Sterelny (1987) is an accessible example. In this study, a major premise is that musical meaning can be discussed against the background assumption of naturalism, and hence we expect a natural explanation for an ontologically accessible phenomenon. The reason that this proviso is important is that the theory about to be put forward is primarily a cognitive one. We are going to be discussing the constitution of musical meaning as a mental phenomenon that is describable in terms of mental states. Therefore, it would be wise to be clear, from the outset, about what is assumed regarding the nature of the mind, and put it on record that these mental states are assumed to have a physical constitution.

A major motivation for this study is grounding the discussion of musical meaning in terms of causal principles and mental states that are open to empirical investigation. Current studies on musical meaning appear to make no fundamental assumptions about the ontological status of meaning, nor as to the status of the mind that plays host to meaning.[8] Of course, the practical problems surrounding the discussion of mental phenomena are daunting, and it would appear unreasonable to expect researchers who are principally interested in music to have a definitive grasp on the issues surrounding the nature of the mind and mental phenomena. In this study, an effort to be clear about such grounding principles is made: a naturalistic conception of the mind and mental phenomena is adopted. That is, the mind and mental phenomena are phenomena which, after adequate conceptual clarification, are subject to empirical validation by scientifically accessible means. The 'empirical validation' proviso is one that must, for the purposes of this study, be considered in principle. Current research on the brain has not yet provided clear-cut answers on the precise mechanisms whereby consciousness and the sensations of subjectivity are achieved, nor as to the fundamentals of communicative meaning as it is neurologically manifested.[9] Nonetheless, there appears to be a growing trend toward naturalistic explanations for mental phenomena, which per definition will be eventually available to scientific validation (see, for example, the interviews in Blackmore, 2006).

[8] Granted, however, that discussions of meaning in the work of Cross and Patel are tacitly congruent with scientific realism regarding the ontology of the phenomena discussed.
[9] For an easily accessible and highly entertaining discussion of consciousness by leading figures from relevant fields, see Blackmore (2006).

There seems to be little reason to think that that which we call the mind is anything other than the brain – namely, that the brain *is* the matter of the mind. This would be a position compatible with (but not necessarily identical to) physicalism, where it is assumed that all mental phenomena are at base physical, and are governed by physical laws that are open to empirical investigation. The reader will note that hard and fast ontological assumptions are not vital to the applicability of this model. However, it might be stated that the favoured ontological assumption is akin to the so-called 'causal closure of physics' position adopted by some philosophers. Let us state, therefore, that all the phenomena described in this study are subject to empirical investigation that is, in principle, attainable through present and future methods of the empirical natural sciences.[10]

Although it is being forwarded that naturalism with regard to the mind should be taken as a given, it is worth (briefly) mentioning one or two issues around the subject. Here we must venture into a field that is also generally subsumed under the banner of analytic philosophy: the philosophy of mind. The philosophy of mind is as old as philosophy itself, being that part of philosophical investigation that enquires as to the nature of thought and consciousness. With a resurgence in interest in the twentieth century because of the field of cognitive science, the philosophy of mind has undergone something of a reawakening as a major topic of interest, becoming an interesting subject which draws researchers from a variety of fields other than philosophy. The mechanisms of thought, issues of perception, the possibility of artificial intelligence, and the fundamental nature of the mind are all topics of interest in modern philosophy of mind. These issues all become intertwined when questions are raised as to the ontological nature of the mind in relation to the physical manifestation of the human brain. Is the mind composed of special subjective 'mind stuff', that is independent from physical processes? This was the view of Descartes, who in 1641 expressed the belief that the mind and the body were composed of different 'substances'. According to Descartes, the body (that is, the brain) is material, but the mind (or '*res cogitans*') was not reducible to physical matter.[11] Unfortunately, this explanation brings about the famous mind-body

[10] Unfortunately, the present author is no expert on modern physics. Nonetheless, the view that in some deep sense there may possibly be some randomness underlying physical phenomenon is the single reason why the term '*causal* closure of physics' has been qualified!

[11] The reasons that Descartes arrived at this conclusion are epitomized in the famous phrase, 'cogito ergo sum' ('I think, therefore I am'). According to this view, the systematic elimination of physical matter does not eliminate the thinking subject itself. Systematic doubt of the existence of the body and all physical reality is possible, but doubting the existence of thinking is impossible, because there must be a subject who is thinking those doubtful thoughts.

problem: how does the mind, which is purportedly non-physical, interact with the material body, when material effect necessarily requires a material cause? While it has been argued that the Cartesian formulation of the problem is flawed (e.g., at length in Dennett, 1991), the problem of how mental phenomena arise from physical phenomena remains at the heart of philosophical discussions about consciousness. Rock art expert David Lewis-Williams puts the stickiness of the problem into perspective when he notes that the mind "cannot be placed on a table and dissected as can a brain" (2002:104).

Despite these problems, the mind and its contents appear to us to be a concrete reality. Our subjective lives are as real to us as the so-called contents of consciousness, our subjective experience of the world. It is this subjective, thinking self and all its experiences that Descartes called the 'mind', Plato called the 'soul', and that philosophers and cognitive scientists attempt to define in relation to the brain. How it is that mechanistic processes in the brain give rise to the experience of what it is like to taste wine and other features of our introspective lives, is the topic of vigorous debate. And, even if we are able to explain consciousness and our subjective experience of the world in physical terms – what Chalmers (1996) refers to as "the hard problem" – we are hard pressed to provide an evolutionary explanation for it.[12] Nonetheless, the present author finds the reasons to assume a non-physical explanation of consciousness less compelling than naturalistic attempts.

What, then, are we to assume when we speak of 'mental states', a major topic in this study? The following explanation should suffice for the purposes of the ensuing discussion. First, note that 'mental states' and consciousness are not equivalent. I am not making an assumption as to the nature of consciousness, although I am going to use phrases such as mental states 'coming to conscious attention'. Second, all activity in the mind must be underwritten by activity in the brain. A change in mental state cannot occur without a change in material arrangement of the brain, and no two individual, differing mental states can have identical arrangements of brain matter underlying them. In other words, a mental event only occurs when there has been a change in the 'matter of the mind', the brain. Any given mental state will have a particular physical constitution in the brain, an arrangement of atoms that is

[12] The problem of the evolutionary rationale behind subjective self-awareness and consciousness has been referred to as "the other hard problem" (Polger & Flanagan, 1999). However, it is likely that knowledge as to the biological constitution of the processes that give rise to consciousness is likely to open a window on its evolution. For example, if consciousness is simply the by-product of complex brains, there would be no evolutionary reason why it exists beyond the selective pressure for more complex brains.

unique to that mental state. Every mental state is therefore uniquely determined by a brain state. It should be noted that I have not said that a mental state is the same thing as a brain state.[13] However, at the very least, a mental state is *determined* by a brain state (that is, the arrangement of matter in the brain). Therefore, when we say that someone is entertaining a particular mental state, the claim is that there is an arrangement of matter in the brain that determines that mental state. For the duration of the study, I will adopt the terminology 'mental state' to describe this state of affairs.

An ambiguity can arise from this nomenclature. When a person looks at a tree and takes notice of the appearance of its branches, which reminds her of clouds, we can describe the exact arrangement of every bit of matter in the brain as this person's 'mental state'. If we were time travel back to before she has looked at the tree branches, and then somehow rearrange all the matter in her brain to match the tree-relevant mental state, she will have the sensation of looking at the tree's branches and being reminded of clouds, even if she has never had this thought before. However, it is also common to speak of there being numerous mental states (plural) in the mind of a person. Thus, 'feeling sad' can be described as a mental state. In this case, we can induce a feeling of sadness by exposing the subject to certain stimuli. We can say that the person entertains several brain states, such as the state of sadness, a state regarding the apprehension of her surroundings, a self-reflective state of the form 'I am feeling sad', and so on. Note that there is an in-built plurality ambiguity to this terminology: in the first case, we can say that a person at a particular time *t* has such-and-such an arrangement of all the matter in the brain, and that this arrangement determines the person's mental state *in toto*. This mental state (in the singular) comprises whatever is the focus of conscious attention, the person's overall mood, physical tokens relating to memory, and so on – everything about mental life, conscious and subconscious, at time *t*. However, in the second case, we can speak of mental states in the plural: conscious attention directed toward a tree's branches, the state of being in a contemplative mood, etc. In this manner, we can say that several individual mental states can be entertained simultaneously. Defined in terms of each other, mental state *in toto* is comprised of various individual mental states. We

[13] I strongly suspect that this is the case, however (that brain matter is in some sense the mental state). I think that the intricacies of the problem of whether mental states simply are brain states is further complicated by the fact that we tend to talk of physical processes as one domain, and mental states as another. Issues surrounding the mind are therefore often approached by use of metaphor, which intuitively inhibits us from simply seeing brain matter arrangement (brain states) as mental states. However, this assumption is not required for application of the model of meaning presented in this study.

will use the plural form of 'mental state' throughout this study, allowing us to claim that a mind is able to entertain several sorts of mental states simultaneously.[14]

1.4 Linguistic modality and linguistic computation

This study is principally about music, and the typically human phenomenon of perceiving musical pieces as being somehow meaningful in a communicative sense. The topic of communicative meaning cannot be successfully discussed without reference to the human communicative medium *par excellence*, language. And to successfully tap our knowledge about language, we must be sure that we are clear about what this unique ability is. It is clear that whatever language is, it can take many forms. One encounters language when the radio or television is turned on, when a book is opened, when we look at a direction sign on the highway, and when we communicate with each other by talking. Therefore, we can note that language manifests itself in several forms, speech and writing being the most obvious. The use of gesture in sign language is another example of a form of language, as sign languages also make use of similar grammatical rules to spoken and written language. The various manifestations of language are its *modalities*. Hence, we have the modality of speech, the modality of writing, and the modality of gesture in the case of sign language.

Since the 1950s, interest in the biological constitution of linguistic behaviour has grown immensely. This has led to an increased awareness of the constituent behaviours comprising language usage. For instance, throughout history, language has principally been used in the spoken form. Doubtlessly, the evolutionary heritage of language as a communicative medium is closely tied to the physiology of our vocal tracts. But does this imply that language is dependent on the physiology of speech? That cannot be the case, as people who are unable to speak are often still able to communicate through writing, which is perfectly intelligible to other language users. Indeed, writing in general is a speech-independent manifestation of language. Likewise, people who were born deaf, and have therefore never heard the sounds of speech, are able to utilize sign language, which depends on similar grammatical rules to

[14] Also note that mental states are not the only sorts of 'objects' in the mind. There are also computational systems for handling data input, which result in various mental states. The computational system 'Language', described in the next section, is one such example.

those underlying other modalities of language (Bickerton, 1996; Botha, 2000).[15] Because we can say that language is implemented in several modalities, we can raise the question of whether there is some sense of 'language' that is independent of those modalities.

Hauser, Chomsky and Fitch (2002) write that any investigations into language, especially those that purport to make claims about evolution, need to take two conceptions of language into consideration. These two conceptions affect a distinction between the communicative system of language considered as a whole, and the computational system in the mind that underlies this form of communication. The former is what the authors call the faculty of language in its broad sense, or FLB. The latter (that is, the cognitive mechanisms underlying the execution of linguistic communication that are common to all its modalities) is the faculty of language in its narrow sense, or FLN (Hauser, Chomsky & Fitch, 2002:1570-1571). The broad sense of the language faculty, FLB, subsumes the narrow sense, FLN. FLB is composed of systems that Hauser, Chomsky and Fitch call the "sensory-motor", "conceptual-intentional", and "computational" systems.[16] We needn't go into too much detail about the precise nature of these systems. It will suffice to note that these are the sorts of systems that allow us to use concepts and a vocabulary, the cognitive equipment for apprehending and producing utterances, as well as the power to operate all the necessary physiological apparatus that is used in various manifestations of language. These are all the things that are necessary for the effective use of (spoken) language as a communicative medium.[17] The FLN comprises the "computational" system alone: the "abstract linguistic computational system... independent of the other systems with which it interacts and interfaces" (p. 1571). Hauser, Chomsky and Fitch believe that the only part of FLB that has no homologue or analogue in any other species is the computational system, or FLN. The linguistic computational system is the only human-specific feature of language as a whole. For these authors, the distinctive feature of FLN is recursion; for Chomsky, recursion is manifested by virtue of syntactic operations of the sort described as 'minimalist' (see Chomsky, 1995).[18]

[15] Superficially, what all these modalities have in common is the use of syntactical rules and arbitrary symbols that are treated semantically. Later in this study, the notions of semantics and syntax will be considered in more detail.
[16] The 'computational' system *is* FLN.
[17] Note that FLB does not include things like memory, or even a living organism that can play host to FLB. These are necessary but not sufficient conditions for FLB.
[18] This sort of distinction can already be seen in early work by Chomsky, such as the distinction between competence and performance (1965).

This is a distinction worth making, especially in light of views expressed by Hauser, Chomsky and Fitch, as well as Bickerton (1996), amounting to the claim the cognitive apparatus underlying the operations of syntax in linguistic behaviour may not have had their origins as organs of syntax implementation, but rather as aspects of cognition evolved for dealing with other circumstances.[19] It is a distinction that we should be aware of. Our primary concern is going to be with language in its communicative sense, as a package. When I wish to indicate the computational system that is language (analogous to FLN), I will use 'Language' (with an upper case 'L'). Linguistic behaviour in general will be indicated by use of a lower case 'l' ('language').

How does music shape up with regard to modality? Should we be speaking of a faculty of music in the broad sense (FMB) and a faculty of music in the narrow case (FMN)? Or, rephrased, is there a computational system for music, and can music be implemented in different modalities? In this study, it will be argued that there is indeed something that can be characterised as a syntactical dimension for music, but its existence as an independent computational system in the mind is a moot point. It may be the case that this system is a combination of various domain-general (as opposed to music-specific) cognitive capacities, and possibly one or two music-specific subsystems. This, along with the reservations about the presence of 'mental objects' such as a deep structure for musical utterances expressed by Lerdahl and Jackendoff (1983:278ff; see also Lerdahl, 2009), will serve as a restraint for speaking of music (or 'Music' in the narrow sense) as a computational system in the mind. Furthermore, the cognitive apparatus underlying musical cognition has a major difference to that of language: music only occurs in one modality, namely sound. It is inconceivable to think that there is a computational system underlying the implementation of musical syntax in different modalities – that is, music in gesture and writing in addition to its regular manifestation in sound. Music is *modality-specific*, and that dedicated modality is sound.[20]

This is an interesting difference between music and language. The cognitive abilities underlying language use (that is, FLN) can be used in various modalities, but the cognitive mechanisms underlying music are dependent on the medium of sound. It can be noted,

[19] The issue of linguistic ontology with regard to evolution has also been taken up Botha (2000).

[20] The case of musical writing, such as a score, is a function of our semantic abilities, as the inscriptions are arbitrary symbols referring to actual music (that is, sound). It will be argued in this study that the syntactical dimension of music is intricately connected to its modality of sound, in that it functions as a means of organizing that specific modality.

however, that linguistic behaviour (in general) does not escape modality-specificity completely. For instance, in ordinary inscription of language by means of writing, there is no commonly used system for mapping the inflection of the voice onto the piece of paper, despite the fact that we regularly use voice inflection to communicate over and above the powers of syntax and semantics. There is no way of indicating the voice inflection that is often used to indicate displeasure, for example. But these features do not reflect what is common to instances of language usage in each of its modalities. What is common to the various modalities in any given language are the syntax and semantics of the language, which are implemented by means of FLN. Non-syntactical, non-semantic arbiters of meaning are specific to the modality that is used to implement the common syntax and semantics. In this study, it will be argued that such non-syntactical, non-semantic features of language are modality-specific, and play a part in the determination of mental states that represent communicative meaning in the mind. The syntax of music differs in that it has a vested interest in the communicative potential of the modality of sound itself: musical syntax exists to organise sound and no other modality.

We have now considered several foundational issues pertinent to this study. We are presently due to turn to a consideration of the concept of musical and linguistic 'meaning', as a property of communicative acts. In doing so, we are going to evaluate the concept of 'communication', and introduce the idea of describing meaning as manifesting itself in the mind by means of mental states.

CHAPTER 2
INTRODUCTION TO COGNITIVE AND AFFECTIVE MENTAL STATES

2.1 Introduction: A model of communication

This study has, as its primary concern, the idea that music somehow boasts 'meaning'. It steps off from the premise that it is fairly common for people who listen to music to attach some sort of meaning to it, or claim that the music somehow complements some situation in an appropriate (or inappropriate) manner. Often, the idea that a piece of music is meaningful is accompanied by questions about what the composer or performer was attempting to *communicate* to the audience. A listener may ask *what* the composer or performer is trying to say with a particular piece of music. What is the reason for the use of such-and-such a chord at this time, instead of a different one? Am I, as a listener, expected to understand what the composer intends with that chord? Has the composer used a musical medium to communicate something that is difficult to verbalise, that could not have been said in ordinary speech?

Music would appear to have a close relation to linguistic communication, as can be easily thought in light of recent work on the biological basis of musical behaviour (e.g. Patel, 2008a). Is music, therefore, a communicative medium? And if so, what does it communicate? Conceptually, this is a tricky question to answer; a question with which thinkers grapple to this day. The sounds that constitute non-programmatic music (that is, music that is not intended to illustrate or accompany some non-musical story or scene) do not have conventional meanings in the manner in which linguistic communities have agreed-upon meanings for words. There exist musical pieces that cannot be said to refer to anything beyond music, in the manner in which a statement such as "Those mushrooms are expensive" refers to the monetary expense of real-life fungi that exist in a world independent of language. And yet, throughout history, there has been an urge amongst humans to employ sound musically in interactive activities, to somehow give an additional sonic voice to their thoughts. The term 'musical expression' has saturated talk about music. But what is expressed, and how does it come to be so viewed?

Cross sees music as a form of communication (1999; 2005; Cross & Woodruff, 2008). However, he notes that the notion of information theory does not entirely capture the

uniqueness of musical activity. Remarking on the generally accepted model of communication proposed by Shannon and Weaver (1949), Cross notes regarding this traditional model that

> a sender makes use of a channel to send information to a receiver, the sender and receiver can be of any type of entity, the channel can be constituted of any medium, and the information that is sent may take any form. In a musical context, one can think of the sender as the performer, the receiver as the listener, the channel as the air and the information transmitted is the sonic patterns that constitute the music.
>
> (Cross, 2005:2)

However, Cross goes on to note that while this model appears to fit well with the modern listening habits of many Westerners, it is not congruent with many other sorts of music. For example, he cites particular Western examples of cases where the role of performer is often not realised (amateur rock bands and recreational choirs whose goal is not public performance). According to the traditional model of Shannon and Weaver, these musical cases do not count as instances of communication. Furthermore, in many non-Western contexts the roles of performer and listener are blurred. It is not immediately obvious that the roles of sender and receiver of information in traditional information theory are fulfilled in cases of (for example) ritualistic music. In such interactive cases, says Cross, "music may be more a medium for participatory interaction where all are equally and simultaneously performers and listeners than a medium for display, for communication of musical information to 'passive' listeners" (Cross, 2005:3). Cross also notes that unlike language, music lacks unambiguous intentionality or 'aboutness' (*ibid.*:5): a non-programmatic musical phrase is not *about something* in the world. Musical utterances do not refer; and if they ever do, those referential meanings are inherently ambiguous to the listener (*ibid.*:8).

With musical statements lacking intentionality and frequently violating accepted informational-communicative norms, it would appear that the case for any clear-cut classification of music as a communicative medium is problematic. However, if we consider communication from a different point of view (that we could possibly call a psychological perspective), musical behaviour may be able to fit the bill as a form of communication despite a lack of unambiguous intentionality and its interactive nature. This would involve viewing the traditional idea that communication involves the transfer of information from the

mind of one person to that of another as a subtle misconception, albeit one that seems generally tolerable. There is no informational entity that, having being appropriated by the communicative medium, ends up in the mind of the listener. Rather, what happens is that some thought in the mind of the utterer is rendered in the communicative medium by virtue of the medium's causal properties. Provided that the listener has the appropriate mental and perceptual apparatus for apprehending that communicative medium, the thought is then represented in the mind of the listener, because of the causal properties of the utterance. Hence, there is a process in the listener's mind: from the perception of the message in the communicative medium to its representation in the mind, which then goes on to constitute appropriate mental states by causal means. Hearing a statement about the quality of a particular lager in English involves the perception of the string of sounds as a statement in English, the application of appropriate linguistic and semantic principles to successfully identify appropriate concepts about lager, and then entertaining a mental state that stands in some relation to the lager your friend is speaking about.

Thus, the ultimate point of communication is the attempt to alter the mental states of others. All communicative media share this important characteristic: the ability to affect the mental states of others. Language and music, our primary concerns, both share this common feature, as do other communicative media – gesture, written language, and sign posts are all media of this sort. Let us examine the case of spoken language, by considering Smith, who wishes to tell his friend Jones something about lager. There is a thought in the mind of Smith; Smith renders this thought in the communicative medium of speech according to the grammatical and semantic rules of the language utilised by himself and Jones. This message in the communicative medium of speech is then apprehended by Smith. The end result of the process is that Jones has Smith's intended mental state regarding lager in his mind. Jones has undoubtedly employed the cognitive apparatus needed to parse the statement grammatically and semantically, but these processes have been spurred into action by features of the message itself: features that have some kind of causal efficacy with regard to communication-relevant cognition.

The picture of communication illustrated above is broad enough to include animal communicative behaviours, if we accept the proviso that in many cases the intent to manipulate the mental states of conspecifics is not necessarily a self-reflective action

requiring a theory of mind.[21] However, it is human communication that interests us at present. With a multiplicity of human behaviours that could conceivably count as communicative media, the reader should be prompted to ask if some are more specialised at certain tasks than others. This certainly seems to be the case. It is very difficult to describe the ingredients of a recipe to someone by simply using spatial gestures. Likewise, it is easier to draw a picture of a vertebrate eye than to use written language alone in explaining what it looks like. But regardless of the aims of individual communicative acts, they all share that similar feature: they result in a change in the mental state of attentive and competent apprehenders. We will return to the question of specialised function shortly, as it is natural to ask what the specialisation of music is. For now, what is important to note is that while the ends achieved by various communicative media differ, the means employed are essentially the same. For communication to work, the mental state of the listener must, to some degree, change as the result of an utterance. Successful communication is when the mental state triggered by the sender's message results in the mental state that the sender *intends* in the mind of the listener.[22] Hence, successfully deceiving someone is a communicative act on this model; as is attempting to relay thoughts about the world honestly.

How do musical utterances shape up in the view of communication discussed above? Let us take an elementary example. A performer plays a piece of music to a listener, with the intention of inducing some sort of state of sadness in the listener's mind. The intended state is not necessarily very precisely realised – after all, musical utterances often seem to have an inbuilt vagueness. However, for the purposes of our example, we may assume that the utterer intends to evoke a state that we can classify as being of the 'sadness' variety. She achieves her aim; the listener experiences a mental state akin to sadness. What has happened in this case is that the performer has sought to elicit a particular mental state in the mind of the

[21] An organism is said to have a theory of mind if it is aware of the existence of other similar minds within the same species. In other words, a theory of mind is present if the organism is aware that conspecifics are able to think in a similar manner. That organism would also believe that other organisms within the same species have a similar conception of its own mental life. For more about the concept of a theory of mind, see Premack and Woodruff (1978). The reason for the proviso of animal communication not requiring a theory of mind is to remove the need for the intent to alter the *mental states* of others, which would not be possible without the concept of conspecifics having mental lives of their own. In cases of animal communication without theory of mind, the communicative act may be purely manipulative. Mental states, however elementary, will be altered by the causal properties of the communicative act in all cases of animal communication. However, the utterer itself would act on the premise of manipulating behaviour rather than manipulating mental states, because without a theory of mind, it would not be aware that others have mental states.

[22] In the case of animal communication, the intent to alter the mental states of others may not be consciously perceived as such. Rather, in animals without a theory of mind, the intent to alter the *mental states* of others may be seen by the animal as an attempt to alter another animal's *behaviour* by means of what we would term a communicative act, in a manipulative fashion.

listener. To do so, the communicative medium has been sound, and the causal properties of the message itself have, as intended by the performer, altered the mental state of the listener in the desired manner. Another example is the intention of an utterer to induce a state of excitement in the listener by means of a vigorous piece of music with (for instance) a fast tempo and complex, driving rhythms. The performance of this music serves as a communicative act, resulting in an excitement-relevant mental state forming in the mind of the listener by virtue of the causal properties of the utterance.

The use of things like gesture, speech, music, and writing to affect the mental states of others is communication. Obviously, communication can be more or less effective in terms of the goals and intentions of the producer of the communicative act, largely due to the choice of the medium of communication and the competence of the parties involved in the production and reception of that communicative act. It is possible to intend to elicit particular thoughts in the mind of a listener, and not succeed. If the lack of success results in zero change in mental state for the listener, we might say that no communication has taken place whatsoever. This can be because of any number of reasons: perhaps the apprehender was out of earshot, or has no means of facilitating the proper perception of the communicative act by not being able to read. If there is a resultant mental state in the mind of the apprehender, but it is not the one that the communicator intended, then we say that the communicator has been misunderstood.[23] Once again, there could be many reasons for such a state of affairs. One such reason is mistaking elements of a communicative medium with other elements in that medium, despite being fully competent (for example, mistaking the referent of one word for that of another; or even having a different referent for a word than the person initiating communication). It is possible, however, for communication to take place despite inadequacies with regard to the utterance itself. This is in virtue of the listener's interpretation of the communicative act as a whole. When the communicative act is incomplete or inadequate in some manner, an inference of what it is that the utterer intended to communicate can be made by the listener. This is pragmatic context.

[23] In the case of music, which is inherently ambiguous, cases of being misunderstood would occur when the listener perceives a piece of music intended to elicit a reaction in some broad class (e.g. 'sadness') as being indicative of some completely alien class (e.g. hearing the same piece as belonging to a 'happy' class). In other words, the communicative ambiguity characteristic of many types of music means that communicative intent cannot be as specific as it is with other media, such as writing.

Before discussing pragmatic context, an exception to the basic communicative rule needs to be pointed out. Sometimes utterances are produced simply because they sound nice to the utterer, or are used by the utterer simply because he or she feels like engaging in musical activity. However, such utterances do not just sound joyful, sad, or profound to a chance listener, but also elicit mental states in a manner indiscernible in type from other communicative acts. While the utterer's intention to communicate was not present, this still does not prevent a process similar to what I am about to describe from happening. A lack of intention to communicate on behalf of the utterer does not prevent a statement in a communicative medium from affecting the mental states of others that receive it. Music played by yourself, for yourself, and for no one else, will have communicative effects on chance, unintended listeners regardless of your own communicative wishes. It is nonetheless true that there exists the fact that while music is predominantly an interactive group activity in the majority of the world's cultures, there are cases where music is performed solely for the benefit of the performer – such as instrumentalists playing for themselves, alone. These non-intentional cases do not, however, change the way in which musical meaning comes about in the minds of chance listeners. Doubtlessly, there is in many such cases a fair amount of extra-musical significance attached to the musical performance by the performer. But such a scenario does not change the fundamental manner in which music changes mental states, even if that change is somehow the result of a reflexive process in the mind of the performer herself.

2.2 Pragmatic context in linguistic communication

At this point, it would be opportune to mention the role of pragmatics in philosophical considerations of meaning and linguistic communication. The following statement by Stalnaker, made with specific reference to linguistic expressions, should set the agenda neatly: "Syntax studies sentences, semantics studies propositions. Pragmatics is the study of linguistic acts and the contexts in which they are performed" (Stalnaker, 1970:275). Whereas syntax concerns the order of elements in a communicative act, and semantics the meaning of those elements and the sentence as a whole, the pragmatic meaning of a statement concerns the effect that the conditions under which the utterance was made have to the overall meaning perceived by the listener. Phrased succinctly, "pragmatics is fundamentally concerned with recovering the thoughts communicators intend to convey by their ostensive acts" (Carston,

2008:340). Let us take the following sentence as an example. Imagine two people, John and Peter, sitting in a room on a summer day. John is sitting next to a closed window. Peter utters the following sentence, directed at John.

(1) It's getting very hot in this room.

John then gets up and opens the window. What did Peter really say to John with (1)? Judging from the meaning of the words alone, Peter's sentence is merely a reflection on the current state of affairs regarding the ambient temperature of the room. On such a literal reading of the meaning of the words of the utterance, the apprehension of sentence (1) would not appear to be the reason that John opened the window. Of course, this is obviously not the case. What happened is that Peter *implied* with (1) was that John should get up and open the window, so that the room could cool down. The literal meaning of the words of the sentence alone reflects the semantics of the sentence, as facilitated by the syntax of the English language. Understood in isolation from the context under which it was uttered, (1) is not a request that John should open the window. However, when considered in context, the obvious implied meaning of (1) is that John should get up and open the window. This second (contextual) interpretation is made in virtue of the pragmatics of the utterance. Even though the words of sentence (1) have retained their semantic meaning, they are being used for another communicative purpose. Put in terms used in this study, the change in the listener's mental state is not just in virtue of the meaning of the words, but also in virtue of the manner in which the utterance is being used in a particular context.

Another well-used example is that of a school pupil who has not been excelling in his schoolwork. On the year-end report card, the teacher writes only the following:

(2) Peter is pleasant and has neat handwriting.

By virtue of the meaning of the words alone, the sentence expresses precisely what it says: Peter is pleasant in the school environment, and his handwriting is legible. These statements can be confirmed as true or false. But what the teacher is really saying is what is conspicuous by its absence: Peter has not got much in the lines of academic promise. This unwritten meaning, understood in the context of report cards, is the pragmatic meaning of the

sentence.[24] As this example and the example before demonstrate, it is possible to say things with a statement that are not encoded in the utterance itself. In these cases, it is not in virtue of the meaning of the words and the syntactical means by which they are combined that a listener (or reader) can understand what the utterer is really trying to say.[25]

At first glance, the fact that the sounds of music do not have conventionally agreed-upon referents in the manner in which words refer to things beyond the confines of language, would seem to pose another problem to those wishing to understand music as a communicative medium. It also makes a clear distinction between semantics (or communicative meaning) and pragmatics in music difficult to draw. This is the view of Swain (1997:77), who reduces musical pragmatics to a matter of stylistic genre. But examples of non-linguistic pragmatics can be formulated. Carston (2008:338-389) provides an example where encoded meaning is kept to an absolute minimum. Two friends, Sue and Bob, are in a pub. Sue is standing near the bar when Bob catches her eye. He waves an empty pint glass in the air, making sure that Sue notices it. Sue takes this visual interchange as a signal for her to order Bob another pint of beer, which she proceeds to do. In this case, there has been no conventionally encoded signal used at all. Waving an empty pint glass is not a conventionally agreed-upon signal meaning 'Order me another beer'. Rather, in the context of the pub, and based on the mutual acquaintance of Bob and Sue, Sue's interpretation of this as a signal for her to order Bob another beer is "derived through a bit of probabilistic reasoning based on general knowledge" (Carston, 2008:339). This is general knowledge of the situation: Sue's interpretation of the gesture is a matter of pragmatics, and has nothing to do with the meaning of symbols.

What is interesting in these three cases is that there is an *intention* to communicate present, even if there is a poverty of relevant encoded meaning and syntax in the communicative medium (the spoken sentence, the writing on the report card, and the gesture, respectively). Pragmatics deals with the recovery of what the utterer, writer, or gesticulator intended to communicate. When framed in terms of communication, it would appear that pragmatics augments the use of any communicative medium – after all, pragmatics is the recovery of

[24] A third example is provided by Aitcheson (1989:269): 'Do you know the time?' If we were to interpret this by the meaning of the words alone, the answer would have to be 'Yes' or 'No' – either you know the time, or you don't. However, that is not what the question is really about. Rather, the utterer wants you to tell her what the time is, not ask whether you know the time. It is the pragmatics of the question – how it is used – that leads to the appropriate reaction. Yet another excellent example can be found in Carston (2008:322-323).
[25] More on pragmatics in debates about linguistic meaning can be found in the essays comprising Grice (1991).

what the utterer intended to *communicate*. Describing how pragmatics in music occurs will involve a description of how the intention behind musical utterances is recovered by the listener in virtue of the context in which the musical utterance is made. It is not, however, written into the music itself, in the same way that the pragmatic meaning of the examples cited above are not intrinsic to the utterances and gestures involved. One would therefore expect to find that musical pragmatics is mostly a matter of the individual contexts in which musical utterances are heard, and depend to a large degree on non-musical factors. For the remainder of the study, we will make note of relevant cases where pragmatics contributes to the musical meaning perceived by the listener.

It should also be noted that information about the mental states of others can still be apprehended without a communicative act having taken place. This sort of 'informational updating' is a fairly elementary way of getting about in the world. You can ascertain that your friend is in a foul mood, based on the manner in which he treats the things in the environment (for example, slamming a door). Such behaviour is a consequence of a mental state, and not an act of communication. To be sure, a large amount of our information about others is gathered in this manner. But this is not the same sort of process as the act of communication. Recall that earlier, it was stated that communicative media could result in a change in mental state in the apprehender. What is important to note is that unintentional use of communicative media will always result in a change in the mental state of an apprehender, provided the apprehender has the requisite competency for the communicative medium in question. Behaviour or other non-communicative factors used in informational updating do not always result in a change of mental state for the apprehender. Observing that my friend is nauseous does not mean that I will automatically believe him to be nervous.

This sort of informational updating should be distinguished from scenarios of pragmatic context, such as those that were described above. With pragmatics, the listener attempts to recover the meaning that the utterer intended to convey. In the examples cited above, Peter intended that John should interpret the statement about the room's ambient temperature as a request to open the window; the schoolteacher intended that positive remarks regarding non-academic matters on an academic report card be interpreted as a negative comments about scholastic ability; and Bob intended that Sue should interpret his gesture as a request for another pint of beer. What is common to these cases is the intention to communicate. However, simply observing that my friend has slammed a door, and therefore inferring that

he is in a bad mood, does not mean that he is trying to communicate to me that he is angry. This deduction is merely made on the basis of non-communicative evidence in the environment. Making an inference that your friend is in pain after an involuntary cry has been issued upon stubbing his toe is another example: here, the involuntary utterance is not made with the intention to communicate to you that he is in pain. The inference that he is in pain is made from a non-communicative, non-intentional (albeit audible) cue. Informational updating occurs when you make a deduction about a conspecific based on non-communicative cues, where there is no intention to communicate present. Pragmatics deals with the recovery of intended meaning that is underspecified in the communicative act. Care should be taken that the two are not confused.

2.3 Music and language as communicative media

What sorts of things can be communicated? In principle, any thought can be communicated, provided a medium that is capable of handling that thought, and a community of at least two interlocutors that are capable of wielding that medium.[26] For example, a principal feature of spoken and written language (indeed, all 'Language' with an upper case 'L') is the ability to communicate propositions. Therefore, all thought capable of being framed in terms of propositions can be expressed by spoken and written language. The only types of mental content that cannot be communicated are those for which no appropriate communicative medium exists. If such mental content indeed exists, we would not be able to represent it in any communicative medium we currently possess.

Questions remain as to the nature of the mental states that are realised in the minds of persons who are on the receiving side of communicative acts. Different sorts of thoughts require different sorts of communicative media, which in turn result in different sorts of mental states. Some states result in beliefs, others in emotion or affect, and others are contributors to a pool of knowledge. The use of the English language, to communicate a proposition about the weather, results in a belief in the mind of the listener with respect to the proposition communicated. The sorts of mental states we are concerned with are those termed cognitive states (which often lead to the formation of beliefs) and affective states (which result in the

[26] By 'thought' I mean any mental state of which a person is consciously aware.

experience of affects). However, before we consider these concepts more closely, we should narrow our discussion down to the two communicative media that interest us the most: language and music.

It has already been mentioned that music is to be thought of as a communicative medium. The goal of communication is the elicitation of desired mental states in the mind of the listener. An obvious question now arises: what nature of thought is communicated with musical utterances? It is clear that music communicates 'content' that is not easily facilitated in a form of spoken language.[27] If musical thoughts were easily represented in speech, we would have to wonder what precisely the point of musical communication would be. In terms of communication, it would not be parsimonious to have one communicative medium that handled all the same sorts of content that could be more easily handled by another medium.

It is possible to imagine that the sort of content that is communicated with musical utterances could be of two varieties. The first is of the sort that could indeed be communicated by some other communicative medium, albeit not quite so easily. For instance, it is possible to speak of music in terms of structure, tonality, and thematic material. These are all concepts that are firmly in the domain of propositional content, and hence are expressible in the various manifestations of natural language (such as speech and writing). We could say that a particular piece has a melody consisting of these notes, and a temporal duration expressible in seconds, and this and that timbre, expressed in terms of the relative amplitudes of its fundamental frequency and partials. But it is far easier to listen to the piece of music in question – all of this linguistically framed information can be actively experienced in a single musical utterance, akin to a sort of demonstrative survey of the piece. It is difficult to imagine any linguistic description of a piece being fine-grained enough to capture every describable element that is heard. Likewise, any linguistic description of the piece is likely to highlight aspects of the music that might not necessarily have come to conscious attention, in the manner in which labels appended to a diagram of the vertebrate eye presents information in manner not identical to actual visual experience of an eye. Listening to a musical piece also grants us access to a different way of representing this propositional content – that is, a language-free description. This is a different way of encountering the music, and brings with

[27] In the ensuing discussion, the term 'content' is used mainly as a matter of convenience. We might describe the content of a mental state representing the proposition 'The grass is green' as simply that statement (i.e. the content of a mental state is that mental state).

it the important aspect of musical experience that cannot be linguistically represented: what it is like to experience the music.

This brings to our attention the second sort of content that we may imagine music being capable of communicating, which is content that cannot be communicated in any other communicative medium. These are, if you will, uniquely musical thoughts, which are inexpressible in any other medium than music. By analogy, we can imagine a similar scenario for gesture, or any other communicative medium. We shall see that this sort of content can, for our purposes, be lumped together with content that is not expressible in the form of propositions. This larger category of content will be called non-propositional content (NPC), of which uniquely musical thoughts are one type.[28]

Spoken language is similar in this regard. It too is capable of representing content that can be represented in another medium. We can describe a vertebrate eye, but the communicative medium of a picture is far easier to use and understand. However, once again, experiencing a description of the eye, experiencing a diagram of the eye, and experiencing the eye *itself* are not equivalent. The second sort of content – that is, content that cannot be represented in any other medium – is also present. Arguably, there are concepts unique to each medium that we would struggle to find any other way of expressing. Temporal concepts are a good linguistic example. How else but with language can we communicate concepts like 'later', 'yesterday', or 'in the future'? It would seem that only linguistic creatures can possess the ability to talk (or even think) about these sorts of concepts. Added to this is the fact that language is tailor-made to express propositions. Therefore, it may make sense to say that language is able to communicate concepts and propositional content, as well as other types of content via performative and pragmatic means.

How far an analogy between music and language can be stretched remains to be seen. For now, we can note that both are communicative media in the sense that utterances serve to affect changes in the mental state of listeners, usually with the aim of the listener experiencing a mental state that the utterer intends. The change in mental state is usually realised by content in the mind of the listener, and that content can either be propositional or

[28] The reader should note that I have stopped short of saying that the communication of such uniquely musical thoughts is the *raison d'être* for music's existence. I think that given current thought about the evolutionary rationale behind musical behaviour, the idea that the need to express uniquely musical thoughts provided the pressure required for selection of musical traits is implausible.

non-propositional, and can sometimes be unique in that it can only be communicated via a particular communicative medium. So far, the analogy between music and language is common to all communicative media. Of course, both music and language (in its speech modality) are forms of aural communication. Both require the perception and production of sound for their effectiveness. This holds the implication that both require some sort of parsing or processing in the mind, in order to differentiate them from other sounds in the environment. The precise nature of this parsing and processing of aural perception with regard to music will be developed later. In anticipation, we can ask whether it is possible that our analogy can be extended in general to the mental processing of music and language, its effect on our mental states, and the sorts of mechanism that allow the implementation of such mental states.

One aspect of comparisons between music and language is the idea that both feature principles for the combination of sonic elements into utterances that are *well-formed*. In language, this set of principles is grammar, whether one speaks of the generative, innate variety, or more culturally-specific grammatical idiosyncrasies. With regard to music, questions of rules and principles for forming acceptable strings of sound are often subsumed by the notion of the musical styles present within a given culture. Thus, we could say that there are principles present within a musical community for combining sounds to create acceptable musical utterances. While a piece may be in violation of certain culturally determined stylistic rules, it may still be comprehensible in terms of deeper, cognitively-based rules. Whether this deeper, universal, and cognitively-based level of principles for music exists seems to me to be beyond question, even though the details of this level and its fundamental nature have not been agreed upon by the broader musicological community. Indeed, in comparison to studies of language, there have been few attempts to investigate this matter.[29] For now, we are able to say that both music and language are aural communicative

[29] Both the matter of syntax and the matter of musical universals are by no means areas of agreement in musicology. This is not to say that attempts at plotting universal features of music haven't been penned, with a notable early example being Erpf (1967). The most obvious musical universal is the fact that, in all recorded societies, the activity of music-making is itself a universal (Peretz, 2006; Blacking, 1995; Nettl, 2005). This is another analogy between music and language, but one that we should be careful of considering remarkable. After all, ritualised preparation of food is another universal in human societies, but we would be wary of attributing this to any special cognitive process. What we should ask is whether there are, within the world-wide music-making community, aspects of the activity that are universal and analogous to language. (We will discount the functional universals of music for now, and focus on the properties of the sonic phenomenon itself.) To do so, we would require an extensive knowledge of linguistic and musicological data, with special attention paid to neuroscientific evidence. For more on the universals of the world's music, see Nettl (2005), especially

media, and that both require the combination of discrete sounds in accordance with rules for their combination into utterances (whatever the fundamental nature of these rules is). The perception of both music and language require the listener to identify the sounds as music or language, and to assess whether the utterance in question is well-formed or not. Provided an intelligible utterance, a particular mental state is elicited in the mind of the listener.

2.4 Mental states as a precondition for musical meaning

At this point, the proverbial elephant in the room is the concept of meaning. In the opening chapter, we established that our primary concern with the concept of meaning is in the realm of communication (as opposed to meaning as significance). With all this talk of mental states, well-formed utterances, and communicative media, it seems reasonable to ask where the meaning of utterances fits into the picture. This is especially pertinent with regard music: what does such and such a musical utterance mean? Is there place for meaning in music? Is this not precisely where the analogy between music and language breaks down?

It is generally thought that, should musical meaning exist, it would differ from linguistic meaning not only in degree but also in kind. For instance, if use of the term 'meaning' is exhausted by a word's referent, it would seem to rule out any possibility of a strict sense of 'meaning' in absolute music (this is the view expressed inn Kivy, 1990). Absolute music does not have words, and musical sounds do not have agreed-upon referents in musical utterances. Of course, it is possible that a musical community could agree on a set of conventions that could be used as arbitrary symbols. We could say that a particular chord progression has a referent *x*, and every time the community hears that chord progression, it is understood that *x* is being referred to. But what we would have is tantamount to a language capable of expressing propositions. Thus, if this were the case, absolute music would simply be used as a language. It would just be another modality for semantics. It would also have the appearance of a contradiction in terms. Non-programmatic music is, per definition, the sort of music that has no extra-musical referents.

---

pp. 42-49. Both the issue of universals and that of the status of well-formed utterances will receive more attention in Ch. 5.

Despite this, it seems to be the case that when most people are confronted with non-programmatic music, they somehow feel that there is something to be grasped or understood, some sort of 'meaning'. In addition, most cultures use different sorts of music for different purposes, depending on which music fits the context best.[30] What is it that is grasped or understood? Scruton (1987) has pursued such a line of reasoning, stating that "[t]he meaning of music is what you understand when you understand it" (p. 169), and that "meaning is the object of understanding" (*ibid.*). This avenue of thought seems to make a fair amount of intuitive sense. After all, we have already stated that music is to be thought of as a communicative medium, and surely this implies that in musical statements, there is something to be understood. This something to be understood is, in Scruton's view, musical meaning.

However, the problem with defining meaning as the object of understanding is that it narrows the applicability of the concept, thus rendering certain commonplace uses of 'meaning' (with regard to music) unaccounted for. The term 'understanding' implies that there is, to some degree, a process in the mind that must take place before meaning is grasped. Upon hearing a musical stimulus, the mind must somehow process the information it has received and represent it in some manner, possibly according to some set of principles that serves an evaluative purpose. Now, I am not denying that such processes occur. Indeed, this was a point in favour of the analogy between music and language: both require the parsing or processing of auditory stimuli, thus differentiating linguistic and musical sounds from others in the sonic environment. But now we are confronted with what, at first, will seem a major disanalogy between the two communicative media. There is a vast array of musical sounds that appear to have effects on us that are not dependant on 'understanding'. Take, for example, a sudden loud chord where one would not expect it. The sudden *fortissimo* chord signalling the beginning of the development section, b. 161, of the first movement of Tchaikovsky's Sixth Symphony, is a very good example. The reaction that the composer intended – one of alarm, no doubt – is not a matter of 'understanding'.[31] Rather, the state of

[30] There seem, in most cultures, to be differences between the stylistic traits of say, music for ritual funeral and music for weddings or music for didactic children's songs. Obviously, there is something about particular sorts of music that affords them for use in certain contexts, something that one must understand in order to use the correct music at the correct time. Regardless of whether these uses are dictated by societal context or not, there are often qualitative differences between different sorts of music that influence a society's attitude with regard to its appropriateness in various situations.

[31] Many other sounds in the environment result in non-voluntary mental (and physical reactions), including sounds accompanying language in the modality of speech.

alarm is achieved by non-syntactical, minimally-processed means. It is a reaction. There is no need for a set of mental processes that identifies the sound as a particular element of an utterance, represents it, and only then sets the cognitive wheels turning that result in a mental state of alarm. Likewise, the association of an increase in volume, thickness of texture, rising melodic material, and rhythmic activity to an increasing state of excitement is also not a matter of processing. Such direct reactions to musical stimuli seem not to be products of a process of 'understanding'.[32]

But how does this detract from the idea that meaning is the object of understanding, and that to understand a piece of music is to know what it means? The reactions described above are, after all, natural. We respond to loud, unexpected, and sudden sounds with alarm, and we treat the source of a general increase in noise as something in an excited state. What has this to do with the perception of music as something meaningful? Consider a piece of music, such as a moderately paced, medium-loud folk tune. Now, consider the same piece of music, only that the last note of each phrase is played with an aggressive attack and at a high volume. Alternatively, imagine the same piece of music with loud, aggressive attacks being used on every note; or, imagine the entire length of the piece comprising a *crescendo* from very soft to very loud. What is the difference between our original folk tune and the versions that are changed with respect to attack and volume? The mental states that result from these various versions are most certainly different. There is a distinct difference in what is communicated in our original folk tune and what is communicated in the altered version, despite the fact that we have made no changes to the notes or the manner in which they are combined. The changes are very much intended to form part of what is communicated to the listener, but these are not changes that result in any need for a type of processing resulting in 'understanding'. They are purely changes of the type that we naturally have reactions to – there is no need to postulate a chain of mental events and representations, a set of evaluative mental processes. But these changes still contribute to what is communicated in a vital way. An aggressive rendering of our folk tune communicates something different than the original version does; the two versions result in different mental states in the mind of the listener. Most would say that the perceived meanings of the two versions are different. If the idea is correct that parameters like the use of loud, aggressive attack result in mental states that are

[32] I am not, however, denying that it is possible for the subject to be artificially conditioned to react differently to aural stimuli, and hence have a reaction to a sound that we would not ordinarily expect. However, such a scenario would involve conditioning to such an extent that the natural reaction to the stimulus is overridden.

not the result of processing (that is, not a process of 'understanding'), then arguably the overall perceived meaning of a piece of music is not exhausted by mental processing.

Musical meaning must, therefore, be made up of at least two sorts of co-existing mental states: *cognitive states*, that are the result of mental processes (such as concepts or data represented in the mind being manipulated by principles of some sort); and *affective states*, that are elicited by directly causal processes (that is, elicited by our hard-wired reactions to sonic stimuli). When a listener expresses the perception of meaning in a musical piece, her report stems from the existence of these two sorts of mental states in the mind: cognitive states that are the products of evaluative mental processes ('understanding'); and affective states, which result from aspects of the musical experience that cause largely involuntary reactions. When the performer attempts to communicate (that is, attempts to change the mental states of the listener), both cognitive and affective states are exploited.

2.5 Causal mechanisms surrounding the elicitation of appropriate mental states

Before continuing, the idea of mental 'computation' must be discussed. Throughout the remainder of the study, mention will be made of processes of mental computation. However, the initial state of the data that is processed should be defined. When someone hears a linguistic or musical utterance (or any sonic stimuli, for that matter), the string of sounds needs to be made amenable to mental operations. In other words, the string of sounds, as apprehended by the sense organs, needs to be rendered in 'mind matter': it needs to be represented or encoded in such a way that the brain can make sense of it as computationally malleable data in its most basic form. Let us refer to input from the sense organs, in a format potentially available to the brain for further processing, as *the stimulus as it is immediately represented in the mind*. Thus, when we speak of features of the musical utterance being in a directly causal relationship with a particular mental state, what is implied is that it is those features immediately represented in the mind that do the causal work. This is a scenario similar to the way in which we might describe the process behind a pinprick that results in a mental state of pain. We would ordinarily say that the pinprick is in a directly causal relationship with the reaction of pain: the pinprick causes the state of pain. However, the resultant mental state is actually in a directly causal relationship with encoded information regarding the pinprick, as it is made available as input to the mind by virtue of the signals of

the nervous system. The entire process can be described as the pinprick causing the nervous system to send a message to the brain, where it is immediately represented or encoded in 'mind matter', and it is this immediate representation that is in a directly causal relationship with the state of pain. This representation is in a computationally malleable state, and has itself already been subject to computation of some basic sort (that is, some sort of computational representation in the nervous system). Therefore, with musical and linguistic utterances, the message is apprehended by the sense organs, and then immediately represented in the mind; and it is this immediate representation that stands in a directly causal relation to the resultant mental state. When we speak of computational processes on this immediately represented data, the reader should note that such data has already undergone a measure of computation in being rendered an immediate, computationally malleable representation of the stimulus in the mind.

Recall that it was stated that the point of communicative acts is to elicit a particular set of mental states in the mind of the listener. Effective and successful communication results in the utterer's successful elicitation of the desired mental states in mind of the intended listener. The argument up to this point has mentioned two specific sorts of mental states: cognitive states and affective states. Before considering the roles of music and language in the achievement of these mental states, some time should be spent exploring the nature of the mechanisms by which these states are achieved.

Assuming a naturalistic view of the workings of the mind that is consistent with the processes describable by the methods of physical science (that is, akin to the causal closure of physics), it should be apparent that all activity in the mind is, at a fundamental level, the result of causal relations in the matter of the brain. Therefore, we should expect that however the apprehension of meaning works, it is bound to be underwritten by causal mechanisms. There are combinations of causal mechanisms that are of such a nature that we can describe them as computational systems. Innate grammatical principles, such as those that are described in psycholinguistics, are examples of such complex causal computational mechanisms. At a fundamental level, things like innate grammatical principles must be sets of causal processes, material causes and material effects. But what is remarkable about causal complexes such as generative grammars and the like is the fact that they are themselves computational systems: a set of causal processes that manipulate data, but are independent of that data. The point of these computational complexes, these sets of causal mechanisms that can be described in

terms of principles, is to manipulate an input and produce an output. For example, sentences need to be processed in terms of grammaticality and the meaning of the individual terms before the meaning of the sentence as a whole can be appreciated and represented in the mind. Thus, the input is the string of sounds constituting an act of speech, as it is immediately represented in the mind, and the output is the general meaning of the sentence.

These computational processes serve various purposes. However, whenever data is processed, we can conclude that the output results in a change in the arrangement of the matter of the brain. In some cases, the outputs will be available to conscious experience. When these outputs do come to conscious attention, we can say that they serve to change mental states. When we hear a spoken utterance, the output of the various computational systems involved will include the general meaning of the sentence *in toto*, semantically irrelevant sonic properties (say, representations of the speaker's accent), pragmatically relevant information, and the actual physical sounds. While the meaning of the sentence is usually immediately available to conscious experience, things like the qualities of the speech sounds and semantically irrelevant sonic details can be brought to conscious attention. In these scenarios, it is fair to say that content is consciously represented in the mind; or, in other words, there is a change in mental state.[33]

Mental states that are the result of systems of computation, as described above, can be termed cognitive states. One feature of such states is the fact that they lead to beliefs about the outside world. For example, the apprehension of speech has, as an ultimate result, a cognitive state that can be used in the formation of beliefs about what has been said. If an utterer says "The British public is very fond of curry", and this sentence has been perceived and processed successfully by an apprehender, that apprehender will be able to form beliefs regarding the representation of the sentence's meaning. The apprehender may form a belief about the sentence as a whole, such as a belief that it is true. By logical means, this can conceivably lead to other beliefs about the culinary preferences of individual members of the British public (for instance, the belief that the British public is fond of at least one type of food). Such beliefs have been arrived at via cognitive mental states.

[33] Note that the earlier distinction between communication and the mere apprehension of information about the speaker applies here. Someone who tells you that your car's headlights are on in a Scottish accent is not also trying to communicate the fact that he is Scottish. That you gain knowledge of the speaker's nationality is a case of picking up information not from the communicative act *per se*, but from the details of its execution (i.e. features of the sonic environment at large).

Note that the very stuff of such computational procedures must be represented in the mind in such a way that allows for the inference of beliefs. Beliefs can only be inferred from propositional content. Let us pause to consider this statement. It is possible for beliefs to be *about* content that cannot be expressed in propositional format, but beliefs cannot be inferred *from* content that cannot be expressed in propositional form. You can believe that you are sad, but you cannot infer things from the state of sadness itself. We have already mentioned the existence of non-propositional content (NPC). It follows that NPC cannot be used to infer beliefs. It also follows that NPC cannot be subject to computational procedures and processes in the mind. If computational procedures result in content that can be represented in such a way as to allow for the formation of beliefs (and other propositional attitudes), and beliefs are propositional, then non-propositional content must be the sort of content that does not allow for beliefs at all. Of course, non-propositional content is perceivably available to conscious experience, so it does constitute mental states.[34] The sort of mental state that arises in the case of NPC is what we called an affective mental state earlier.

Two questions now confront us. The first concerns the nature of the mechanism by means of which the apprehender of a communicative act, such as the listener of music, arrives at an affective mental state. The second question is whether a situation analogous to cognitive states leading to the formation of beliefs occurs with regard to affective states. (Or: if cognitive states lead to beliefs and other propositional attitudes, do affective states lead to something analogous?)

The first question: It has already been mentioned that affective states are the result of causal processes, in a more direct manner than the evaluative propositional-computational procedures underlying the achievement of cognitive states. Now we will consider this idea further. The conclusion we will arrive at is that NPC is realised in affective mental states, and that with affective mental states the stimulus, as it is immediately represented in the mind, stands in a *directly causal relation* to the resultant mental state. Thus, with a communicative act that results in non-propositional content in the mind of the apprehender, the means by which that act is delivered stands in a directly causal relation to an affective mental state. Let us suppose that someone listens to a piece of music. There are plenty of features about the

[34] The contents of subjective conscious experience ('qualia') also constitute non-propositional content.

piece of music that can indeed be represented linguistically. However, recall that there are features of musical experience that are like reactions to sounds in general (b. 161 of the first movement of Tchaikovsky's Sixth Symphony was an example). We had originally said that these were aspects of music where there was no process of 'understanding' involved. There was no manner of process in the mind that receives a loud, sudden sound as an input, and then manipulates it into an evaluated representation by some complex set of principles underwritten by causal processes in the mind, resulting in a mental state that then leads to an appropriate reaction to the sound upon becoming the focus of conscious attention. We noted that such a chain of events violates parsimony, and that these sorts of reactions could be better described as hard-wired reactions. The mental state that results from experiencing such sounds is not the result of evaluative, propositional-computational processes.

Assuming the causal closure of physics, we also noted that the activity of the mind is fundamentally underwritten by causal relations. Hence, the evaluative computational procedures leading to cognitive mental states can be described in terms of complexes of causal relations between the various bits of matter in the brain. Because all activity in the mind is, at a fundamental level, underwritten by causal relations between bits of brain matter, it would follow that the mechanism by which affective mental states are realised must also be causal in nature. The difference, however, between the causality of cognitive states and that of affective states is firstly a matter of complexity, and secondly a matter of function. The computational processes for realising the cognitive states described above are complex systems of causal relations, assembled in a complex manner in order to fulfil a particular function. This sort of mental organ exists independently of the data that it handles, and the data it handles does not impinge on the structure of the organ. Computational systems like this exist for the sake of computation. Computing and processing an input to produce a manipulated output, available to other systems or to conscious experience, is the purpose of such an arrangement of matter in the brain. Evaluative computational systems are not only complex entities, they are also there to function as a producer of input for other processes in the brain, and ultimately to realise cognitive states. Input is acted upon by a causally underwritten computational system. Now, causal mechanisms in the brain are either acting upon data, or they are reacting to data. Computational processes exist to act upon data, but what of the data that is *reacted to* by causal processes in the mind? Such a relation – a directly causal relation between data and the matter of the mind which it encounters – is what characterises the realisation of affective mental states. It is the by-passing of complex

evaluative computational systems by certain types of immediately represented data that results in affective states.

To summarise the answer to the first question (what is the nature of the mechanism by which affective mental states are realised?): affective mental states are realised when the communicative act contributes an immediately represented input that is not acted upon by complex sets of evaluative, propositional-computational processes, but rather acts upon the matter of the brain in a directly causal manner. Such a situation will result in the realisation of an affective mental state. The nature of this mental state will provide the answer to the second question posed above: is a situation analogous to the inference of beliefs from cognitive mental states present with regard to affective mental states, given that inference only occurs when content is propositional in nature?

I would like to argue that there is a roughly analogous situation with regard to affective brain states. For lack of a better term, we can say that affective states lead to the apprehension of *affects*. (I have purposely chosen to avoid the term 'emotions' in favour of 'affects', as 'emotion' has too narrow a connotation.) Take, for example, a communicative act that results in an affective state (the Tchaikovsky example is a good one). After experiencing such a stimulus, we should expect a state to be present that has been arrived at via directly causal processes. With the Tchaikovsky example (a loud, sudden orchestral chord after a long period of quiet respite that was ended with an extremely soft downward line in the bassoon), one could conceivably imagine a state of alarm. Such a state has been affected by directly causal processes, and was not arrived at by a set of computational procedures similar to those described in relation to cognitive states. This state leads to awareness of an affect: in our example, the affect that we often name 'alarm'. How would this sort of affect differ from propositional attitudes, which result from cognitive brain states? Most noticeable is that the state is non-propositional, and therefore, no inferences can be made from it. What is there to make an inference about? Can one state of alarm provide the means by which we can make inference about other states of alarm? Of course, we can name the state, but this sort of naming simply points to the particular affect, and does not explain the characteristics of that affect (or class of affects).

Thus, affective states lead to the apprehension of affects, and neither affects nor affective states are propositional in character. It follows that affects cannot be verbalised (although

they can be named, and those names can be used in propositions). They are, in essence, beyond verbal description. Affects are *ineffable*, being cases of non-propositional content. The qualities of affects like alarm, or possibly even a sense of foreboding triggered by a sequence of musical sounds, are not expressible in terms of propositions, and are not the products of cognitive states and evaluative computational procedures.

2.6 The argument so far

Let us take stock of the present state of the argument. Communicative acts have the principal purpose of eliciting a desired mental state in the mind of the listener. If the desired mental state is achieved in the mind of the listener, we can say that the utterer's communicative act has been successful. In principle, anything that can be thought can be communicated, provided a communicative medium capable of articulating those thoughts and a community of at least two people that understand that medium. Some sorts of thoughts are only expressible in a certain medium – these thoughts are unique. Others are translatable in other communicative media. The principal reason for choosing one communicative medium over its translation in another is ease of use, although what is capable of being communicated in a particular medium is also a major motivating factor.

Meaning, it was claimed, is constituted in the mind by at least two co-existing mental states: cognitive states and affective states. Cognitive states are the result of processes in the mind that are in a deep sense underwritten by causal relations. These processes are complex, and serve an evaluative, propositional-computational function. This means that they exist independently of the data that they compute. The outputs of these evaluative computational processes result in cognitive states, which are propositional in character. Cognitive states lead to propositional attitudes (such as beliefs, hopes, etc. – *that*-clauses). Cognitive states are instances of propositional content in the mind. Affective states, on the other hand, are the result of directly causal processes in the mind. Here, the immediately represented input stands in a directly causal relation to the resultant state; there is no intermediary system of causal relations constituting further computational processes which act upon the input. Affective states result in the awareness of affects. Affective states are non-propositional. This means that they cannot of themselves be part of computational procedures, and nothing can be

inferred from them. They are also not expressible in the form of propositions: they are in effect ineffable with respect to our linguistic abilities.

Our next task will be to continue the analogy between music and language, in an attempt to see how far it can be stretched. It has already been noted that music and language, both being communicative media, serve to elicit changes in the mental states of listeners. These states are either the product of directly causal or causally-underwritten, propositional-computational procedures. Together, these two sorts of states account for the perceived 'meaning' of a communicative act, as they are forms of content in the mind of the listener. In short, a listener's perception of the meaning of an utterance is reducible to content represented by cognitive and affective states. When we talk about a listener's perception of the meaning of a piece of music or a linguistic act, we are in fact talking about a composite of these two types of mental state. The analogy between music and language stems from their common classification as communicative acts. Hence both elicit cognitive and affective mental states in the mind of listeners.

CHAPTER 3
STRUCTURE, SYNTAX, SEMANTICS, AND CONTEXTUAL MEANING IN MUSIC AND LANGUAGE

3.1 Introduction

In the previous chapter, we noted that music and language can both be considered as communicative media, and that the meaning of an utterance in a communicative medium is realised in the apprehender's mind by two sorts of co-existing mental states. These mental states differ in terms of the mechanisms by which they are achieved. It was suggested that complex, evaluative and propositional-computational processes acting on data, but remaining independent of that data, were the sort of processing mechanisms that led to what was termed cognitive states. In such cases, data underwent the operations of a computational system in the mind. Cognitive mental states are of the sort that allow for the formation of propositional attitudes, such as beliefs, which essentially constitute cognitive states themselves. It can also be noted that cognitive mental states are truth-apt, and can only process the type of data that is represented in the mind in propositional format. On the other hand, affective mental states were the result of the immediately represented communicative stimulus not being subjected to computational processing, but rather eliciting the mental state in a directly casual fashion. Such mental states were not propositional in character, nor could truth-aptness be assigned to them. Rather, the sort of content comprising them was termed non-propositional content (NPC). It was also suggested that affective mental states are termed affects – that is, states of alarm and so on.

The difference between the two mechanisms can be easily seen when one considers that every process in the mind is underwritten at a deep level by some sort of causal mechanism. This includes the constitution of the propositional-computational processes that result in cognitive mental states. While these processes are in reality made up of a complex of fundamentally casual mechanisms, they are independent of the data that is computed. Hence, the input data is not directly causing the resultant mental state; rather, it is happening in conjunction with an otherwise encapsulated computational procedure. On the other hand, with affective states, the input data is in a directly causal relationship with the resultant mental state. The immediately represented input data resulting in a cognitive state is

subjected to processing by computationally complex systems in the mind. Affective states, being constituted by non-propositional content, are not subject to evaluative, propositional-computational procedures. It follows, per definition, that NPC cannot be expressed in the communicative medium of natural language.

The following question now arises: if language and music are both communicative media, and communicative media elicit mental states in this manner, how do the two differ? At this point, it would seem that what has been asserted is that the elicitation of mental states in music and language do not differ in kind. How then, do they differ? In order to answer this question, it is necessary to once again take up the analogy between music and language. Where the analogy holds, we will be able to make assumptions about the nature of the two communicative media. Where the analogy begins to fail, we will have an indication of the differences between the two behaviours.

3.2 Structure in music and language

There are many compelling reasons to pursue an analogy between music and language. The literature on the neurological relations between music and language cognition is growing (e.g. Patel, 2008a; Peretz, 2006). However, what we wish to understand is whether this analogy can be pursued on a conceptual level, and specifically in the terms described in the previous chapter. To pursue this conceptual examination more closely, we will have to begin to distinguish between what parts of music and language contribute to the two sorts of co-existing mental states. One aspect is the way in which musical and linguistic utterances are actually put together. The keen reader will note that in the following discussion of structure in music and language, I have assumed that there are constituent elements of which musical and linguistic utterances are composed. In language, these constituent elements are formatives and words, combined into phrases. With regard to music, I will later develop the view that there is a grouping principle, largely dependent on perceptual constraints, which demarcates similar structural units. For now, we will assume that music too is composed of discrete elements (such as individual notes of varying duration) that are combined to form musical utterances.

Music and language, it is often claimed, share certain structural similarities. Without yet examining the precise nature of these structural similarities, we should consider examining the idea that organised structure exists at all. This is hardly a controversial issue with regard to language: some orderings of English words make sense to English speakers, and others do not. Analogously, it would appear that in music, some orderings of notes seem to make musical sense, and others do not. It would seem therefore that there are acceptable and unacceptable ways to combine constituent elements in both communicative media. Violating these criteria of acceptability results in music or language that 'doesn't make sense' to speakers and listeners that are fluent in the language and music in question.[35] This implies that the sounds of communicative media are perceived and processed structurally, and violation of structural principles results in a lack of comprehension with regard to the utterance. In music, it would appear that there are many ways to combine sounds in making comprehensible and acceptable phrases. However, anyone who has listened to a toddler bashing randomly at a piano keyboard should be aware that there are also sequences of sounds that make no musical sense whatsoever.

The fact that the order of constituent elements is significant to the comprehensibility of communicative acts allows us to say that both music and language have a *syntactic* dimension (on a particularly weak understanding of the term 'syntax'). In other words, the ordering of the constituent elements that make up a linguistic or a musical utterance has an impact on the comprehensibility of that utterance (that is, our ability to make sense of it). At this point, we are not going to make any assumptions about the precise nature of this syntactical dimension, such as what precisely counts as well-formed with regard to syntax in music and language. All that needs to be noted for present purposes is that, with both musical and linguistic utterances, there are examples of well-formed and ill-formed statements. An ill-formed sentence is incomprehensible: we may be able to find out what the words mean, but if they are haphazardly combined, we will be unable to make any judgement regarding the meaning of the sentence as a whole.

With an ill-formed utterance, syntax makes absolutely no contribution to our identification of a string of sounds as a communicative act. Take, as an example, a sentence in a natural

[35] Another way to illustrate this is to imagine yourself as a blind anthropologist, encountering an undocumented society for the first time. It is conceivable that you would be able to distinguish musical from other (say) ritual sounds, because of the presence of basic syntactical principles in musical statements, even if you are not familiar with the intricacies of the musical style. These principles are the focus of the fifth and sixth chapters.

language that is simply a random jumble of words, with no ordering whatsoever. Unless it is well-formed by complete chance, the listener will have to depend on other cues to identify the string of sounds as an unsuccessful communicative act. Recognisable words with agreed-upon meanings are used in sentences like "Man dog bite", and their combination in the act of elocution accounts for us identifying this as a communicative act of some kind. But the syntactical ordering of the words is ill-formed, and does not contribute to communication in any way.[36] Analogously, a random sequence of sounds that does not subscribe to syntactical principles will also not be recognisable as music. To identify a sequence of sounds as an attempt at a musical communicative act, we usually would have to rely on other cues. The example of a toddler bashing away at a keyboard is a paradigm case – the only thing that would link this sequence of sounds to the idea of a musical statement is the fact that the sounds are coming from a well-known musical instrument.[37] Ignoring intentionality, the need for a performer, and the aims of communication for the moment, a random sequence of sounds in nature can only be heard as a musical statement if they happen to satisfy syntactical demands in some way (for example, a regular pulse, identifiable pitches in a well-formed order, and so on).

One may note that there is a grey area with regard to the well-formedness of statements: sentences that are, strictly speaking, ill-formed, but we are still able to correct in our heads. An example is simply omitting words from a sentence, such as "Charlie fell down hole", which is missing an article such as 'a' or 'the'. Here, we are still able to figure out meaning to a large extent, namely, that Charlie has fallen down some or other hole. However, with "Man dog bite", we cannot work out much beyond the fact that the sentence is about a man, a dog, and either a noun or a verb in the form of 'bite', if there is no pragmatic context to help us. We cannot identify a subject and a verb phrase, and hence we cannot attribute an action to either the dog or the man. We will return to this grey area, because something analogous is to be found in the syntax of musical statements: statements that are not quite well-formed, but recognisable in terms of musical syntax.

[36] Unless, of course, there is an appropriate pragmatic context. For example, the ill-formed utterance 'Man dog bite' is perfectly intelligible if there happens to be a man with a canine-inflicted injury present. In such a case, the listener will deduce from the context that what was intended was a communicative utterance that amounts to 'The man was bitten by the dog'.

[37] Another example is musical bottles. In that case, there are orderings of hitting bottles that are syntactical and understood as music, and others that are just heard as noise.

There is another feature of syntax that is notable at this point. Syntax in language cannot be perceived if the apprehender does not know the meaning of any of the words in the sentence. It is impossible to judge a sentence using French words as well-formed or not unless you are familiar with those French words. The French sentence may be of subject-predicate form, with a noun phrase and a verb phrase, but this knowledge is of little value to someone confronted with words whose meanings are completely alien. What words constitute the verb phrase, and which ones the subject? It is impossible to tell without some tentative knowledge of the lexicon of the language.[38] We will return to this feature when we consider semantics below, as it holds interesting implications for our discussion of music.

We have not ventured deeply into the differences between music and language in terms of the precise nature of their structuring principles, but we have stated firmly the premise that both language and music boast some sort of syntax (i.e., that the ordering of constituent elements makes a contribution to the comprehensibility of the utterance). We shall investigate syntax in more detail in later chapters, when our focus shifts to the details of musical syntax. We will also then briefly consider two claims that are often made, especially in evolutionary musicology: that the structuring of music is hierarchical, and that the structuring of music is recursive. Both of these claims deserve mention, but cannot be handled unless we are on firm conceptual ground. They are, however, of importance. Recursion is earmarked by some as the single distinguishing feature of linguistic computation (Hauser, Chomsky & Fitch, 2002), and it would be of interest to see if and how it could be realised in musical utterances. For now, we will proceed to consider the issue of semantics.

[38] The reader should note that I have stopped short of saying that we must know the meaning of words in order to perceive the syntactical features of an utterance. Indeed, it is often the case that the syntax gives us the very meaning of a word. For example, the meaning of 'bite' is determined by the syntax in the following two examples: (a) "That's a nasty bite" and (b) "He can bite you". In (a), 'bite' functions as a noun (an injury); and in (b), 'bite' functions as a verb. Note also that the meaning of the sentence as a whole contributes to cases that would otherwise be equivocation: (c) "I need to draw some money, so I am going to the bank" and (d) "I want to catch a fish, so I am going to the bank". Nonetheless, there are ambiguous sentences, such as (e) "Bill saw the boy with a telescope" (example from Gillon, 2008:376). Here, it is syntactically unclear whether Bill used a telescope to see the boy, or whether Bill saw a boy who had a telescope in his possession. In such cases, we are likely to rely on the context in which the sentence was uttered to interpret the intended meaning.

3.3 Semantics in music and language

As with our discussion of syntax, we will have to forestall complexity in the following discussion of semantics. What we need at this stage is to establish what semantics is, and how it could be discussed in terms of our current project. Semantics, in colloquial terms, deals with meaning, usually understood as a quality or characteristic of natural and artificial languages. However, we are trying to explain 'meaning' in music, so a similar definition will not serve our purposes. Looking beyond matters of definition, we can note that semantics is studied by a group of people who are of considerable interest to our endeavours: analytic philosophers of language. Let us consider the term 'semantics' as it pertains to this field. Semantics is a discipline within analytic philosophy that is primarily concerned with studying the conditions under which sentences in a natural or artificial language can be said to be true or false.[39] Stated in simpler terms, semantics can be said to be the philosophical discipline that studies how meaning is facilitated by a language.

Consider the following definition, formulated by logician Alfred Tarski: "*Semantics* is a discipline which, speaking loosely, *deals with certain relations between expressions of a language and the objects* (or 'states of affairs') *'referred to' by those expressions*" (Tarski, 1944:435; his emphasis). The 'relations' posited by Tarski can be expressed as the truth or falsity of those expressions. An example will suffice to provide the flavour of such philosophical pursuits. Consider an English expression like this:

(1) My cat is sitting on my bed.

We say that this expression is *true* if and only if there is a state of affairs in the world, such that the cat that belongs to me is sitting on the bed in question. This state of affairs is the expression's *truth condition*. If there is no cat, or it is not on my bed, or it is standing on my bed, or any other violation of the truth conditions, we say that the expression is *false*. If it were somehow logically impossible that my cat could sit on my bed, then we would either deem the expression false, or claim that it is somehow *meaningless* (this is a fine distinction, that depends largely on one's theoretical position regarding the manner in which terms refer to reality).

[39] Note that in linguistics, semantics is often construed as the study of the meaning of words or phrases, whereas in analytic philosophy it deals with the conditions under which propositions can be said to be true or false.

One should note, therefore, that in terms of semantics, expressions in two different languages that have identical truth conditions mean the same thing. Thus, (1) and the Afrikaans expression "My kat sit op my bed" mean the same thing, and the same state of affairs will render both true or false. (At any particular given time, the one cannot be true and the other false.) Therefore, it makes sense to say that whatever those two sentences mean, it is the meaning of the sentences that is true or false (if semantics alone is considered as the determinant of meaning). We call this meaning, this hypothetical 'thing', that is true or false, a *proposition*. Therefore, we can recast Tarski's definition as a standard definition of 'semantics' (SDS):

> (SDS): Semantics is a discipline which deals with certain relations between propositions and the objects or states of affairs referred to by those propositions.

This immediately leads us to the question of whether music can be said to boast semantic qualities, when 'semantics' is this narrowly construed. The reader is likely to recall that earlier, we noted that if various musical statements had agreed upon meanings, accepted within a community of listeners, and that it was used to refer to things, then music would simply be a sort of language (with a small 'l'). When music is not used in this manner, it does not handle propositions and therefore cannot bear truth.[40] Here, I am in agreement with philosopher of music Peter Kivy (1990). Music, if it is used in a manner where sounds have been arbitrarily agreed upon by a linguistic community to refer to things beyond music, is acting as a language (albeit an ineffectual and clumsy one). Music can easily be made to mean something *by convention*, just like the sounds in language mean by convention. But music does not always function like this. Therefore, it would seem that semantics, defined as in (SDS), is not a necessary feature of meaningful music *per se*. With a bit of effort and the co-operation of a community of listeners and utterers, music can be made to behave like language, in that it is a system of organised sound acting semantically. But this is a feature stemming from the fact that both music and language have been found to be communicative media. The fact that there is a thing called language that handles propositions, and a thing called music that does not, suggests that at a level of analysis less fundamental than that of communicative media, there is a point of disanalogy between music and language. Language

[40] Indeed, when one considers a view such as Stalnaker (1970), where semantics is defined as the study of propositions, it becomes clear that music has no semantics if it is unable to express propositions.

and music, at their extremes, differ in that the one handles propositional content and the other does not.[41] We will see later that this difference can be explained in terms of the mechanisms by which music and language elicit mental states.

Let us take stock. Semantics is defined according to (SDS), whereby it is seen as dealing with the relations between propositions and relevant states of affairs in the world. These relations are expressed as truth conditions. Language, with its ability to express propositions, has semantic characteristics. However, music can only have similar semantic characteristics if we treat it as a language, by assigning referents to sounds that have been agreed upon by a linguistic community. This is a matter of definition. While we have said that both music and language can be thought of as communicative media affecting the mental states of competent listeners, the case of semantics and language's position as the handler of propositions seems to mark a major difference between the two communicative behaviours. When we treat music like a language, by assigning conventional meanings to the elements that comprise it, we are more inclined to think of it as a sort of language, because we understand language as being the communicative medium that has semantics.[42] What is the importance of this observation, that the case of music having or not having semantics is a matter of definition? The effect of these remarks is that the case of music not having semantics is the outcome of what would otherwise simply be a verbal dispute: that arguments where music has semantics must define semantics as something broader, usually an all-encompassing conception of 'meaning'; or erroneously treat non-programmatic music as if it has its own semantic dimension. Conventional symbolism does not, per definition, function in non-programmatic music. It is meaning in non-programmatic music that we are trying to explain.

3.4 Non-communicative contextual 'meaning'

A task that now remains for this introductory discussion of mental states and musical syntax is to account for an idea that can be found in much recent musicology, namely that the meaning of music is largely determined by the non-musical context in which it is heard

[41] It is worth mentioning at this point that it is common practice for philosophers to talk of 'meaning' as if it were exhausted by semantics (truth conditions, reference, etc.). However, we have already adopted a much broader conception of meaning, as the amalgamation of mental states in the mind of a listener.

[42] A good example that immediately springs to mind is the use of leitmotivs in Wagner's operas. An instance of a particular leitmotiv signalled to the listener that a particular character or concept should be brought to mind, by virtue of the fact that the leitmotiv was earlier directly associated with that character or concept.

(whether that context is social, political, gender-determined, economic, and so on). The pursuit of the contextual 'meanings' of music has been the focus of much mainstream musicology since the mid-1980s, and the term 'meaning' has been subjected to much misuse. Often, semantics is conflated with this sort of meaning. Alastair Williams, in a reflective book on the state of musicology, manages to discuss 'semantics' without ever defining it, tacitly equating it to some sort of conception of meaning that is largely determined by social "discourses" surrounding the music (Williams, 2001:42-47).[43] Similar is Lawrence's Kramer's *Musical Meaning: Toward a Critical History* (2002), which suggests that meaning in music is a socially determined affair, but fails to distinguish between social significance of music for people and the matter of meaning in terms of the intelligibility of musical utterances. Jean-Jacques Nattiez's widely read *Music and Discourse: Toward a Semiology of Music* (1990), in its discussion of music as a system of 'signs', seems to equally blur any distinction between significance attributed to a musical piece by a listener and meaning as a property of the experience of musical utterances themselves.[44] Nattiez, despite discussing and acknowledging a difference between meaning that is intrinsic to music and meaning that is extrinsic, further confuses issues by talking of "musical semantics" (e.g. 1990:9 fn).

Seeing as we have already accepted the phenomenon of musical meaning as the experience of an amalgamation of mental states present in the mind of the listener in virtue of the causal properties of musical utterances, it would appear that non-musical cases have been ruled out. Our principal interest here is precisely those cases where it is the properties of the musical utterances themselves that determine the experience of music as meaningful. However, this second idea of 'meaning', as related to the extra-musical significance that a musical piece has for the listener, merits brief discussion. For this purpose, it is worthwhile distinguishing *intra-musical (intrinsic) meaning* and *contextual (extrinsic) meaning* – but the reader should at all

[43] In fact, the section Williams labels 'Semantics' is actually a discussion of narratives in discourse and music, and does not even mention the word semantics in any place but its title. With the opening sentence of the first chapter, Williams claims that a major concern in his book "is that music is embedded in discourses and surrounded by ideas that contribute to its meaning" (2001:1). He never goes on to define 'meaning' and distinguish the idea of meaning as anything other than social (for example, as a property of communicative acts).

[44] See, for example, his definition on p. 9 (Nattiez's emphasis): "*An object of any kind takes on meaning for an individual apprehending that object, as soon as that individual places the object in relation to areas of his lived experience – that is, in relation to a collection of other objects that belong to his or her experience of the world.*"

times be aware that the use of the word 'meaning' in both concepts does not equivocate them.[45]

We are going to accept the 'context' as all the factors that appear to affect the experience of music as meaningful for the listener, that are not derived from the properties of the music itself. The contextual dimension is far more unprincipled and variable than the sort of intra-musical properties we have thus far been discussing. Broadly construed, there are two sorts of contextual determinants of meaning: (a) those that, despite not being intrinsic to the music in question, are dependant on features of the other determinants of meaning (such as structure and syntactical arrangement); and (b) those that are only very loosely attached to the music, and are not really dependant on the nature of its physical composition at all. 'Attach' is the operative word for case (b): the contextual determinant of this latter variety is the sort of significance that is implied when people *associate* pieces of music with specific events in life, events that are themselves extra-musical. For example, a particular song on the radio (lets say, a particularly jovial one) might evoke feelings of nostalgia, because the song was current in bygone youth. Clearly, the structural and syntactic nature of the song would not impinge on the feeling of nostalgia that accompanies a hearing of it.[46] However, the case where the contextual determinant is somehow dependant on (but not intrinsic to or derived from) other music-specific determinants of meaning, is of interest. A similar view of meaning has been referred to as homological fit (Cook, 2001a:172) or affordance (Clarke, 2005), and it deserves more attention. However, we should first clarify the general notion of contextual determinants of meaning or significance a little further.

We have already considered the fact that both music and language have structural features that contribute to our perception and comprehension of utterances, and are partly responsible for that perception of mental states we call 'meaning'. By altering these fundamental structural features, we alter those mental states, and hence the perception of meaning.[47] A cursory consideration of this fact about intra-musical meaning should put the nature of contextual determinants of meaning into perspective. The contextual determinants are not intrinsic to the music itself. Rather, these determinants can be divided into two groups, those

[45] Koopman and Davies (2001:262) point out instances of a similar distinction being made between intra-musical (or "formal") meaning and extra-musical meaning. Examples, in respective order, are embodied and designative meaning (Leonard Meyer), intrinsic and extrinsic referring (Jean-Jacques Nattiez), introversive and extroversive semiosis (Roman Jakobson), and endosemantics and exosemantics (Wilfried Nöth).
[46] Indeed, this is none other than the 'Darling, it's our tune' syndrome (Mithen, 2005:18).
[47] That is, intra-musical meaning.

that depend on the structure of the music, and those that don't – (a) and (b) above. Alluded to above were those determinants that are not dependent on the music, case (b). What of those that are? Consider the following example, which we shall use to illustrate both sorts of contextual determinants. Operas by Wagner are not particularly well-received by the Israeli authorities, because of the composer's anti-Semitic sympathies. However, there is nothing intrinsic to the musical structure of Wagner's works that is anti-Semitic. Indeed, we can imagine a fictitious case where Wagner had written in the style of Mozart, and a lot of his music had become equally famous. In such a hypothetical situation, there would still be an association with the anti-Semitic attitudes of the composer. Wagner's music, even if it were written in the style of Mozart, would still attract the ire of the Israeli authorities. This is plainly not because of the music itself, but rather because of the attitudes of its author. This sort of contextual meaning, equivalent to case (b), is associative, and it is essentially *extra-musical*.

However, it is sometimes the case that a piece of music seems to lend itself well to particular interpretations, and not others. Successful interpretations would not work as well if the manner in which the music is constructed had been different. In other words, this sort of meaning is dependent for its validity on properties of the music in question. This is akin to case (a), discussed above. A particularly famous example is that of Susan McClary, who likened the arrival of the recapitulation in the first movement of Beethoven's Ninth Symphony to the "murderous rage of a rapist incapable of attaining release" (McClary, 1991; as quoted in Treitler, 1993:36). Setting the controversy that this statement generated aside, it would be difficult to imagine that McClary could have arrived at this colourful interpretation had the recapitulation begun with a quiet statement of the opening theme, instead of crescendo and loud bravado. McClary's 'reading' of this part of the music is not intrinsic to the music itself, but it is compatible with it. Her view would not have been tenable had Beethoven instead chosen to reduce the excitement at this point. Note that the murderous rage of a rapist is not encoded in the music. It is something that McClary brings to the music herself, and the music is of such a nature that her contribution has in some quarters been taken seriously.

Clarke (2005) has suggested that in cases like this, the music itself *affords* particular interpretations, but at the same time does not afford others. This is similar to the way in which a chair affords sitting to a tired person. In a bar fight, the same chair would afford the

function of a weapon, even if it was not designed to perform this duty. But the chair could never afford making coffee, or printing a thesis. These sorts of things are beyond the logical possibilities that the chair affords. Similarly, the Beethoven example that provoked McClary's much-discussed interpretation could afford several others: calamity, bravura, the menacing return of a diabolical main character in a plot, and so on. But it could not afford the image of an idyllic pastoral scene.

Note that a similar scenario exists for language as well. Both spoken and written language can be seen as having an extra-linguistic dimension of meaning over and above that governed by the grammar and the meaning of the words. For example, a sentence out of *Mein Kampf* has, for 21st century readers, the added dimension of being written by Hitler, which results in associative extra-linguistic meanings. However, this added meaning – or better yet, this added significance – is independent of the grammar and the meaning of the words concerned (provided, of course, that we have chosen a fairly politically neutral sentence from the book). Likewise, singular words or phrases can have connotations for a particular community, that they do not have for other communities. The fact that these connotations are peculiar to one community and not another shows that they are not intrinsic to the nature of the utterance itself.

Therefore, contextual determinants of musical and linguistic meaning are of two varieties: those that are afforded by the utterance (McClary's Beethoven), and those that are totally independent of the utterance (nostalgia for old pop tunes). Both these determinants are not intrinsic to the utterance itself: two listeners in different environments and contexts will have different ideas of this sort of meaning. Contextual determinants of meaning are generally extra-communicative, although it is easy to imagine them being employed pragmatically in the service of communication. One could say that the utterance does not of itself elicit this sort of thing: rather, music either provides the means (by affordance), or the listener simply imposes a general meaning of his or her own, irrespective of the structure of the utterance (extra-musical).[48]

[48] The use of the word 'meaning' in this sentence illustrates the sort of equivocation that takes place in the vast majority of discussions about musical meaning.

3.5 Some more comments about syntax

We have noted that the contextual determinants of meaning are divorced from the utterance itself. Communicative acts in and of themselves only result in the elicitation of cognitive and affective states in the mind of the listener with determinants that are intrinsic to those acts. It should thus be clear that the sorts of significance-relevant scenarios that have been listed as contextual are not intrinsic to the communicative act, although they may depend on an act's structure and be used to communicate pragmatically. Turning our attention to music, it would make sense to say that if music has a set of 'design features' as a communicative medium, we would expect that the contextual determinants of meaning may take advantage of them. But these design features were not tailored for contextual determinants *per se*. The way music is put together may enable views such as that of McClary, but it would not make sense to say that the facilitation of such interpretations is the reason that music is put together the way that it is. Rather, it makes more sense to say that the structure of music is anchored in its nature as a sort of non-semantic communicative medium. Contextual determinants are no doubt exploited by composers – that is, after all, what musical quotation and stylistic parody is all about. But it would be difficult to substantiate the claim that written into the architecture of music are elements of the music's context. There is no reflection of society in the way in which utterances in natural language are put together – syntax does not reflect context. Rather, the way that they are put together reflects cognitive and processing demands.

Music and language both have a syntactical dimension, and this dimension makes a difference to whether we perceive an utterance as well-formed or not. Earlier, two issues were mentioned which we must now take up. First, it was mentioned that there is a grey area with regard to the perception of well- and ill-formed utterances in music and language. It seems to be that some utterances are perfect in terms of grammaticality or the rules of a musical style, while others are less perfect but still comprehensible as instances of music or spoken language. On the other end of the spectrum, there are a large proportion of utterances that are perceived as ill-formed, and if it wasn't for our knowledge of the meaning of the elements making up the utterance, or other cues such as observation, we would not recognise the utterance as a communicative act. Second, we must address the idea that in language we need to know the meaning of (most of) the words in a sentence before we can grasp the syntax, whereas in music we have no semantics to speak of with which to help in the identification of syntactical cues. Let us take these issues up in order.

The idea that there is a grey area between well- and ill-formed utterances in communicative media is best illustrated with musical statements. If music has a syntax, it would seem that a phrase (thought of as functionally equivalent to a sentence in language) would be able to be so ill-formed and random in construction that one would not recognise it as a statement of a musical nature. For instance, a random, unprincipled sequence of notes played on a piano by a toddler lacks syntax, and is hence not perceived as music. Rather, the sounds produced on such an occasion are 'just notes', and the only real reason that we hear them as 'notes' is because the toddler happens to be eliciting sounds from an object that is often used to make music. This is an example of an ill-formed musical utterance. On the other hand (and once again, without intricate details), if our toddler played 'Happy Birthday', this would be an example of a well-formed musical statement. However, if our toddler played every third note of 'Happy Birthday' a semitone flat, then most would probably recognise it as 'Happy Birthday' – just with wrong notes. The wrong-note version is an example of an utterance inhabiting the grey area between well- and ill-formed utterances. Note that if we had an example of 'Happy Birthday' in which every note was incorrect, or if we had an example where so many notes differed from the tune that we would not be able to recognise it as a musical piece, the utterance would be judged as ill-formed. Of course, in such cases, the passage could be played with different notes, and then simply end up as a different piece that is well-formed. But if the order of notes was not syntactical, it would be no different from the toddler bashing away at the keys of the piano.

The reader should note at this point that the old philosopher's chestnut, 'How many wrong notes need to be played before a piece by Bach is no longer a piece by Bach?', is in part a perceptual issue. The importance of this problem to music philosophy at large is because of its links to the question of the ontology of musical works. The problem and the question of whether it is at all relevant will be taken up in the final chapter of this study. What should rather be asked (from a perceptual point of view) is the question regarding when sounds cease to be music at all. Setting aside the issue of style in the musical genre or culture in question, as soon as we are not able to perceive a piece of music as well-formed to any degree, we are faced with an ill-formed utterance that is unrecognisable and non-syntactical. In such cases a piece of music ceases *to be perceived* as a piece of music. The perception of music is of importance here, and one should carefully distinguish between questions about the ontology of a 'musical piece' as a philosophical question, and the perception of music as a cognitive-

perceptual issue. Few would deny the term 'music' to John Cage's compositions, but it is clear that there is no perception of musical utterances going on in a piece like *Imaginary Landscape No. 3*.

This brings up an interesting question that has been raised with regard to accounts of musical syntax such as those of Lerdahl and Jackendoff (1983) and Raffman (1993): if we are willing to accept 'Happy Birthday' as a well-formed utterance, and 'Casta Diva' from Bellini's *Norma* as one too, why are we more inclined to say that Bellini's aria is more profound or moving? If one accepts that this is a valid question, then any sense of profundity cannot be a function of well-formedness, as both pieces are well-formed.[49] Roger Scruton, in *The Aesthetics of Music* (1997), has made precisely this argument, specifically against Raffman's account of ineffable musical experiences. Scruton's argument was geared against Raffman's computational explanation of ineffability. Ineffability, claims Scruton, "is a mark of the aesthetic experience, in all its higher forms" (1997:200). In short, Scruton believes that the aesthetic experience of music cannot be reduced to a computational description such as Raffman's (with profundity and ineffability being part of aesthetic experiences).[50] I will argue otherwise when we pick up on the question of ineffability later. However, for now it should be noted that well-formedness is not the source of any sense of profundity attributed to pieces of music. But I will argue that well-formedness is a contributor to the mental states that afterwards account for our sense of aesthetic experience as a *qualitative* experience.

The second major problem that we had to address was the fact that, in language, one has to understand the majority of the words being used in an utterance before the syntax can be properly discerned. Granted, there are other factors, such as perceptual cues and phrasing, that allow us to distinguish the various elements that go into a linguistic utterance. However, a large part of grammar's contribution to the meaning of a linguistic utterance can only be discerned if the majority of the words of which the sentence is comprised are understood. A problem arises when we try to consider musical utterances in an analogous fashion. There

[49] Note that 'Happy Birthday' and 'Casta Diva' are both well-formed (or, at least it must be assumed so for the purposes of our argument). One is not *more* well-formed than the other. Therefore, one cannot say that profundity (or related senses of awe) is related to the degree of well-formedness.

[50] Note that Scruton's use of 'aesthetic' is a name for what he considers artful, and not a term covering the experience of beauty (or other features) in general. Scruton would not say that listening to Nirvana allows for any experience of 'the aesthetic'. However, I use 'aesthetic' in a broader sense, allowing discussion of 'the aesthetics of Nirvana', for instance. In other words, my use of 'aesthetic' is not prescriptive, but rather descriptive.

appear to be no ‘words’ to understand, so as to make better sense of the syntactical elements of the phrase. This is a major disanalogy between music and language. To answer this question fully, we will have to consider the constituent elements of musical utterances that are put together in a syntactical fashion, and we will have to consider how these differ from the words that make up phrases in natural languages. It is this disanalogy that is a key to understanding why language and music differ in terms of communication.

CHAPTER 4
CAUSAL MECHANISMS AND AFFECTIVE STATES IN MUSIC AND LANGUAGE

4.1 Introduction

As we are implored by cognitive musicologists Patrik Juslin and John Sloboda (2001b:3), any science of music must explain the emotional responses that are often closely tied to musical experience. That music can have a powerful grip on our emotional lives seems to be a fact that is so often pointed out that it has become a cliché. Music, we are often told, is a language of emotion and a way of expressing ourselves. The expression of emotion or affect is the sort of thing that prompts us to engage in musical and other artistic activities. Leonard Meyer has put forward a theory where the entire concept of musical meaning is principally predicated on emotion (Meyer, 1956). Cooke (1959) has claimed that music is a language of emotion. Arguments about music's emotional-expressive powers have also been at the heart of philosophical arguments such as those of Davies (2003) and Kivy (1990; 2002). There have been numerous investigations into the psychological nature of the emotions elicited by music (e.g. Juslin & Laukka, 2003, is an excellent and comprehensive example; as is Juslin & Sloboda, 2001a).

The idea has also crept unexamined into other sorts of musical scholarship, including evolutionary musicology. Brown (2000) contends simply that music's meaning is emotion, and that music and its postulated evolutionary precursor are communicative media tailored to express emotion. On a firmer conceptual and empirical footing, Dissanayake (2008) suggests that we may find music's origins in infant-directed speech, the original purpose of which was the regulation of emotion in infants. Mithen (2005) suggests that the rationale behind musical behaviour is emotional expression. However, the idea that emotional expression is universally intrinsic to musical activities is not entirely supported by ethnographic evidence (Cross & Woodruff, 2008). Take, for example, Venda children's songs. The famous study by Blacking (1967, but see also his 1976) contains numerous examples of songs whose functional existence is not predicated on emotional expression, but rather on the didactic virtues of the text and the actions that accompany performance.

Care must be exercised when considering an 'emotional' account of music. Precisely what do these authors mean with the term 'emotion'? In general, one can take for granted that they mean mental states such as those of happiness, sadness, anger, and so on. But what of apprehension? Or fear? Or a feeling of elation that accompanies the triumphant ending of Beethoven's Fifth Symphony? Or the emotional flavour of blues music? The generally aggressive timbre of the heavy metal genre? And what of the sense of profundity that often accompanies musical experiences? It would appear that such 'emotions' are also of the sort that Juslin and Sloboda would like a science of music to explain. But is 'emotion' not a misleading term for all these diverse mental states? Is a sense of profundity or a sense of foreboding distinguishable from baser, more general emotional states such as sadness and happiness? And is it a foregone conclusion that musical experience will always have this sort of emotional dimension? I will be assuming that such distinctions can indeed be made. However, for the purposes of this study, it would not appear immediately necessary that this distinction needs to be painstakingly delineated. After all, a sense of foreboding and the emotions elicited by a piece sensed to be joyful are both examples of affective states. A sense of foreboding and a state of happiness are both non-propositional mental states. Hence the use of the term *affect*: this term seems broad enough to encompass emotions such as happiness and sadness; as well as general 'moods' that may be induced or attributed to musical pieces and the more complex sorts of feelings sometimes experienced when listening to music, such as a sense of profundity or foreboding. Any mental state that is non-propositional is an affective state, and has all of the essential qualities that have thus far been attributed to such states.

The explanation that we are about to encounter has already been mentioned with regard to another complexity of the experience of music – the idea that something *ineffable* is expressed by musical statements (Raffman, 1988; 1983; Scruton, 1997). (Often, the matter regarding emotion and music is further complicated when authors such as Raffman speak of 'ineffable *feelings*'.) It is with the directly causal processes that result in affective states in the mind that we will look for answers to our questions about ineffability in music. We will need to expand and further elucidate our theory of affective states, and the causal means by which they come into existence. This explanation will undercut arguments such as those of Davies and Kivy, which seek to explain where in the musical process 'emotion' is to be accounted for. Instead, we will consider music as a communicative medium that can be used to induce certain types of mental states in the minds of listeners; types of mental states that include, but

are not limited to, the nebulous concept 'emotion' and ineffable feelings. Put crudely, musical statements, like other communicative acts, thus serve the purpose of changing the mental states of others. There is no need to pursue a line of argumentation suggesting that music 'expresses' emotional content. A musical communicative act changes the mental state of competent listeners. The resultant mental state may indeed be a state that we can classify as an emotion, but it needn't be. It could be a sense of foreboding or profundity; it could be relatively inefficacious in terms of affect. Indeed, in the theory that I am about to describe, we will see that musical activity does not have hard-and-fast principles regarding whether this or that mental state is achieved by listeners, or the degree to which such states are affective. However, it is with the two types of mental states that we described earlier that communicative acts in general do their business. Of the afore-mentioned mental states, cognitive states and affective states, it is the latter to which we will now turn. In doing so, I would like to conceptually account for the purported emotional and ineffable content of musical experience.

4.2 Affective states and language

The reader will recall that we considered two sorts of mental states, cognitive and affective, which constitute the mental manifestation of what is commonly called 'meaning' in communicative media like music and language. The point of musical and linguistic statements is to induce the listener to form cognitive and affective states that are at least analogous to those that the utterer intends. Cognitive states were said to be the result of evaluative, propositional-computational processes in the mind. Examples of cognitive states are *that*- states: X believes *that* Y, X fears *that* Y, X hopes *that* Y. (These sorts of states are often referred to as propositional attitudes.) Affective states, on the other hand, were said to be the result of directly causal processes, where the stimulus (the communicative act represented in the mind) is in a directly causal relationship to the resultant mental states. Affective states are states of fear, alarm, happiness, and so on. It is now our task to consider the idea of affective states more closely. A number of peculiar properties were ascribed to affective states earlier. For instance, it was said that affects are the sorts of things that can be named, but beyond that, are essentially ineffable and beyond verbal description. This is a direct result of the fact that affective states are non-propositional in character. Affective states are also not parts of processes of deductive computation, as such computation can only

proceed on mental content that is propositional in character. The affect 'fear' cannot be part of a logical operation; however, the cognitive state of 'fearing that Y' may well be.

Emotional mental states are states of mind such as sadness, anger, happiness, and so on, and are examples of typical affective states. We can satisfy ourselves with a fairly loose definition of 'emotion', because such mental states are subsumed by the class of states we have been calling affective states (and hence, once again, the use of the term 'affect' instead of 'emotion'). The vast majority of emotions that we feel inclined to accept as instances of 'emotions', as well as a whole host of private feelings for which we do not necessarily have names, are cases of affective mental states. Affective states also account for other sorts of feelings that we may not easily think of as emotions – such as benevolent feelings, feelings associated with suspicion, and so on. Also included are other sorts of states, such as the state of fright (which many would hesitate to classify as an emotion), as well as those sorts of occasionally non-intentional emotions referred to as 'moods'.[51] Affective states also subsume the sort of feeling that accompanies the alleged expression of loftier ideals in music and literature; the sorts of objects of expression that prompt people to describe works as virtuous or profound. The state of profundity is an affective state (but note that the belief *that* something is profound is a cognitive state). So, when John Blacking suggests that certain passages in the third movement of Mahler's Tenth Symphony "express something about life and death and man's struggle for fulfilment and spiritual peace" (1976:61), he is attempting to describe a sort of affective state that he experiences. Note that Blacking says that the music expresses "something" about those topics – a case in point regarding the ineffability of such feelings.

Our account of musical experience is going to deal with affective states to a large degree. Indeed, the idea will be forwarded that, in the absence of semantics, music monopolizes affective states. However, both cognitive and affective states are common to the apprehension of all communicative and expressive media. To illustrate the difference between cognitive and affective states more clearly, and to prepare for the discussion that follows, let us turn to spoken language for a few more examples. Imagine a situation where a researcher angrily

[51] Some possible further clarification: moods are taken as non-intentional because they appear to have no object of intentionality. When you wake up in the morning in a foul mood, there is often no particular thing with which you are angry. Of course, some event may have triggered your mood, and in that case, it may be more accurate to say that your general attitude is being influenced by your attitude toward that particular event. Drawing such distinctions is a prime example of how difficult handling concepts such as 'emotion' and 'mood' can be, as well as serving as an example of why that broader concept of 'affect' was chosen over 'emotion'.

reprimands a student for speaking loudly in the library. Let us say that this researcher is particularly short-tempered, and shouts angrily and at the top of his voice: "You are not allowed to talk in the library!" Ignoring for the moment the requisite perceptual processing of phonemes and similar elements needed for understanding speech, we could say that the comprehending listener (the student) will simultaneously experience the two sorts of afore-mentioned mental states: cognitive and affective states. Now, the data that can be computationally encoded in a propositional manner – the semantic meaning of the sentence, and truth-apt statements regarding the manner of performance – will form the cognitive state. However, one could easily see, because of the manner of the sentence's delivery, that the researcher did not merely want to point out that talking in the library is forbidden. He actually wanted to tell the student to keep quiet. On a particularly strict semantics, it is obvious that this is not the linguistic meaning of the sentence in isolation. The part of the utterance that is conveying this information – the fact that the researcher would like the student to keep quiet – is the manner in which the utterance is delivered in combination with the pragmatic context. (We will ignore pragmatic context for now.) It is the mental representation of this aspect of the meaning that leads to the affective state. The affective state resulting from the researcher's impassioned utterance is the listener's reaction to the researcher's anger, as expressed in his tone of voice. Anger (or something analogous to it, such as extreme irritation) was precisely what the researcher had wished to let the student know about. If the researcher is very intimidating, the reactionary affective state in the mind of the listener may be one of fear, or at least something very similar. This state, the state of fear, is the affective state. Note that the student's belief that the researcher is angry, the belief that she is in a state of fear, and any pragmatic deductions, are all cognitive states distinct from the affective state of fear itself. The actual state of fear, and not propositional attitudes, are the affective states (that is, the affects).

Let us take another example. Imagine an utterance in a language that the listener does not understand. For instance, our ill-tempered researcher may be French, and makes a semantically identical utterance to the one above, only in his native tongue. The listener, the student, does not understand the French language. In such a case, the content of the cognitive state is greatly diminished. However, there will be a certain degree of intent that could be apprehended by the student by virtue of the manner of performance of the utterance. The researcher would usually count on the semantic meaning of the sentence being conveyed along with the additional determinant of meaning in the form of the angry manner in which

the sentence is uttered. The researcher intends that both aspects of his utterance are apprehended by the noisy student. Because of her inadequacy with regard to French, the student will not be able to reach a cognitive state with regard to the semantic meaning of the utterance. However, she will be able to ascertain that the utterance is being delivered in a way that conveys the researcher's anger about something. The listener will once again experience the affective state of fear, as a result of the apprehension of the manner of delivery of the utterance by the researcher. Hence, the affective state of fear is still achieved.[52] The listener will be able to entertain cognitive states such as the belief that the researcher is angry or that the researcher is using a particular tone of voice to emphasise what he is saying. However, the listener will not be able to have a cognitive state about the semantic content of the utterance, or any other cognitive state that is logically entailed by the semantics of the utterance, simply because she does not understand French.

As a third example, consider an artificial language which has semantics, but seems to lack an effective mode of delivering statements featuring an affective component. (That is, a case where the modality of a statement of the language has no consequence for affect.) This is an interesting example. Let us postulate a language, called Binary Language (BL), and say that it has a dual purpose as a programming language for computers, and as a language in which a community of speakers is competent. Now, the language can state the most mundane things about objects, such as this sentence, translated into English: "File 3 has been deleted." This sentence can be a statement of BL, written in a notation unique to BL, on a piece of paper. Granted that the community holds no strange conventions regarding the meaning of writing on pieces of paper, one could say that the performative aspect of this utterance has been minimized drastically, if not eliminated. There is no convention in writing that allows the writer to express the sentence of BL angrily or joyously: the modality of this statement results in no affects. However, the semantic content is plain to see. Cognitive states arising from this statement of BL include beliefs about the fate of File 3, beliefs that are logical implications of this statement (such as the belief that there is one less file present, or that File 3 will no longer be available for copying), and so on. However, the situation surrounding affective states is a bit more subtle. Without any inflection to add to the meaning (such as the sentence being uttered with an angry inflection), there is no affective state invoked by virtue of the

[52] This is not necessarily an identical state of fear to the one that is elicited when she is able to understand the utterance (when the researcher speaks English). But an affective state of that sort will be experienced nonetheless, as a case of pure reaction to the performative qualities of the French utterance.

utterance's performance. Of course, it may be the case that for the apprehender of the sentence, the deletion of File 3 is disastrous, and evokes a state of fear. It may be the case that File 3 is required to prove the apprehender's innocence in a court of law. In such a case, the statement results in fear of punishment. Then we have an affective state of fear, as well as a cognitive state, namely, fear that one will be punished. But this affective state of fear was not caused by the performative aspect or manner of delivery of the utterance – it is caused by the semantic meaning of the utterance. If the utterance was written on paper and the reader did not understand BL, there would be no stimulus that would allow for the triggering of the state of fear. The state of fear for a competent apprehender is caused by the semantics of the statement, and is not a directly causal relation between the delivery of the utterance and the affective state. This is an important and subtle point, and one to which we will return, as it is the source of an interesting distinction between the causal processes underlying music and language.

In the examples above, one would be able (in principle) to list the cognitive states (such as beliefs and other propositional attitudes) and the affective states (such as the state of fear) that are present in the mind of a particular listener. The last example (the artificial language, BL) is a case of affective states not being directly caused by the performance of utterances. Otherwise, the sorts of cognitive states elicited by utterances of natural language include both the semantics of the sentence (provided the listener is competent), and propositional attitudes about the semantics, pragmatics and performance of the utterance. The manner of performance is also closely tied to affective states. In terms of performance, a difference in articulation results in a different determinant of the meaning of the utterance as it is understood by the listener, when meaning is understood as being underwritten by these mental states.

4.3 Affective states and music

I have contended that the apprehension of both musical and linguistic utterances results in cognitive and affective states in the mind of a competent listener. We have considered three examples of how this occurs with regard to spoken language. Now, we turn to examples of musical utterances. In the course of the ensuing discussion, the reader will note that in the apprehension of musical utterances, affective states are far more pronounced than in

linguistic utterances. This is interesting, because it seems to conceptually support the idea that in some respects music and language differ in degree and not in kind when it comes to the sorts of mental states that are elicited in the minds of listeners. As with linguistic utterances, we should take care to carefully describe the processes that occur in the following examples.

When listening to a piece of music, there are several points that we may wish to remark upon. For instance, we may wish to comment on the way that the music is put together, such as the syntactical structure of various phrases, or how the music is constructed in traditional analytic terms. We may also remark on the pitches, the tonality, or other aspects of the musical work. We may also be led to philosophical discussions of whether the work is separate from its performance, or whether each performance is a manifestation of the work (e.g. Dodd, 2000; Ingarden, 1986). We will lay such concerns aside, and rather focus on perceptual issues. This means that we are concerned with individual performances. We may think that the performance of a particular piece was poor, or unremarkable, or particularly moving. We may also say, in the case of music that is familiar, that the piece has been performed or interpreted in a particular manner. Interpretation is usually exercised in terms of the technical execution of the piece, the way in which it was played. For example, a fast piece may be played 'aggressively' or 'neatly' or 'with relentless excitement', and so on. It is these characteristics of performance, as well as the performance of the piece itself, that concern us when considering the elicitation of mental states from pieces of music (as opposed to the qualities of the 'work', when it is seen as being in a different ontological category than performances).

Let us start by considering a piece of music, which is performed by an utterer. The utterer has an audience, the apprehender or listener. The aim of the utterer is to alter the mental states of the listener, usually to mental states that the utterer intends. Let us say that the utterer wishes to elicit a state of sadness in the listener, and let us say he achieves this goal by means of his musical performance. How can we describe this state of affairs? The listener's mental states may be as follows. If the utterer has been successful, we would expect the listener's mind to entertain a state of the sadness variety. This state of sadness is an affective state, and should not be confused with associated cognitive states. Cognitive states that are pertinent to this example will be states such as the belief that the music expresses sadness, or the belief that the utterance has indeed resulted in an affect, namely sadness, and anything propositional that can be said about the performance and the music itself. For now, consider simply the

affective state of sadness, present in the mind of the listener. Let us postulate that there are various varieties of 'sad' states, and that this one, in particular, is unique to musical expression (at least with regard to the state of knowledge of the utterer). In other words, this particular state of sadness, to the knowledge of the utterer, can only be communicated in a musical idiom, and by this musical utterance in particular. Only particular musical utterances will result in this sort of affective state of sadness.[53] The utterer is successful in his communicative act if this particular state of sadness (or at least a similar one) is elicited in the mind of the listener, and we can accept that he is for the purposes of this example.[54]

The affective state of sadness is in a directly causal relationship with the communicative stimulus (or, more accurately, the representation of the stimulus in the mind of the listener). The reader will remember that this is part of our general definition of affective states: affective states are directly caused by the actual act of elocution, as it is immediately represented in the mind. This affective state of sadness has not been achieved by means of deductive propositional-computational processes on represented data. The affect of sadness – the state in the mind of the listener – is not a type of content in the mind that is propositional in any way. It is purely a mental state, and it cannot be part of any logical operation, much in the same way that a chair cannot be part of a logical or computational operation. (Rather, propositions concerning or about chairs are involved in logical operations, but not chairs themselves.)

At this point, the obvious elephant in the room is semantic cognitive states. Musical utterances are not semantic, so we would immediately expect that there are no semantic components to purely musical utterances, and hence no cognitive states that concern semantics. This is a point of difference between musical and linguistic utterances: musical utterances cannot elicit semantic cognitive states. But, as a result of the musical utterance, there are many other sorts of cognitive states that can arise. For instance, a belief that the utterance expresses sadness is such a cognitive state. Another sort of cognitive state involves the means by which the utterance is delivered: the listener may believe that the way that the

[53] That is not to imply that musical statements only elicit music-specific affects.
[54] There is a certain amount of leeway to be expected in such cases. It is unlikely that the medium of music is able to reliably trigger the exact mental states that the utterer intends. However, the states triggered in a successful musical utterance will be of a similar class to those intended (for example, an intended state of sadness). This sort of ambiguity is discussed at length in Cross (2005).

musical statement is uttered is intended to communicate something. This latter instance is an inference from pragmatics.

Take a familiar tune as another example: a folk song, such as 'Greensleeves'. In a performance of this folk song, there will be cognitive and affective states elicited in the mind of the listener. Examples of cognitive states include propositional attitudes concerning the affective state itself, beliefs about the ultimate goal of the musical statement, beliefs about structural aspects of the music, and so on. The affective state will be a result of aspects of the performance whose immediate representation in the mind is in a directly causal relation with that state of mind, with no processes of computation or logical manipulation of propositional data. Changing aspects of the musical performance that are in a direct causal relationship with mental states will therefore result in a change in those mental states. In other words, if the directly causal parameters of the musical utterance are changed, different affective states will arise in the mind of the listener. If our folk song is performed in a particularly aggressive manner, corresponding affective states will be experienced by the listener.[55]

The manner in which a particular folk song's performance changes the affective state of a listener is an illustration of only one parameter of musical utterances that acts in a directly causal manner. The music itself (independent of its performance) can be constructed in such a way that it acts in a directly causal manner, thus resulting in affective mental states. The reader will recall the example of the sudden loud chord in the first movement of Tchaikovsky's Sixth Symphony. This is an example of a structural event in a musical piece that results in an affective mental state. An event such as this loud orchestral chord would elicit something like a state of alarm. The musical stimulus, as it is immediately represented in the mind, is in a directly causal relationship to the affective state of alarm. Other examples may include tension building figures such as the rhythmic sequences opening Mozart's *Rondo alla turca* (from his sonata, K. 331) and Schubert's *Der Erlkönig*, *accelerandos* and *ritenuetos* in general, and climaxes based on volume and increasingly thick orchestration (such as Ravel's *Bolero*). These latter types of causal parameters are the sorts that are built

[55] It should be mentioned that a particularly pertinent example of this is the heavy metal genre known as thrash metal, which was popular in the 1980s, especially in the United States of America. Harsh vocal timbre, high tempos, and aggressive distortion were combined with extremely quick successions of repeated notes to create an aggressive, driving sound. The sound of especially the electric guitar parts in this music is a prime example of affects being achieved by timbral and rhythmic means – almost, some would say, a direct translation of aggression into sound.

into musical figures, and they are the types of cases where we would expect to find a degree of universality.

The actual composed structural features of musical utterances, and the manner in which an utterance is delivered, are two examples of aspects of the musical piece (as it is immediately represented in the mind) causally affecting brain states. These two aspects add to our perception that a piece of music is meaningful, by contributing to an amalgamation of affective and cognitive states in the mind. I have mentioned that music differs from language in that a far greater emphasis is placed on the elicitation of affective states, and that music does not have cognitive states that are in any way semantic. Music's lack of semantics has already been accounted for, so now we should investigate more closely as to what parts of musical utterances serve as the directly causal triggers for affective states.

4.4 More detail about parts of musical utterances resulting in affective states

In summary, the process of arriving at an affective state in the mind of a listener can be described as follows. The musical stimulus is apprehended by the sense organs of the listener, after which it is encoded in such a manner as to be as direct a representation as possible in the mind of the listener. This is the immediate representation of the musical utterance in the mind of the listener, 'translated' into a sort of mentalese that allows for the remainder of the process. This part of the process is not necessarily conscious, for the representation will literally be unprocessed by any other perception-related mental organs. It is merely the musical act represented in 'mind matter'. From here, there is some sort of elementary separation of those aspects of the representation that result in cognitive states and those that result in affective states. Those resulting in cognitive states must be able, in principle, to undergo propositional-computational procedures. In other words, the parts of the immediate representation that lead to cognitive states must be able to be manipulated by computational complexes in the mind. Such aspects of the immediate representation of the musical utterance are propositional in character. Note that not all aspects of the representation that are propositional in nature will necessarily become available to consciousness. Aspects such as the frequencies of individual overtones are apprehended and represented within the limits of the capabilities of our mental and physical apparatus. However, while the relative strengths of these overtones can be expressed in terms of propositions, they are not perceived as

overtones, but rather as a specific timbre. The aspects of the immediately represented musical utterance that do not fit the computational and propositional requirements are those that stand in a directly causal relationship with the mental states we label as affective, or otherwise have no relevant effect on our conscious mental states.

This chapter began with a brief discussion of the idea that music is inextricably bound to our emotional lives. In the course of this discussion, further refinement of the term 'emotion' was abandoned in favour of the use of the word 'affect', to better account for the wide range of feelings often associated with musical experience. This simultaneously accounts for the idea that music is not merely about expressing emotion, but may also have a wider range of influence in inducing mental states that seem to be beyond verbal explication. In such a manner, we can account for the alleged ineffability of some musical experiences – the idea that music is used to communicate ideas that are somehow beyond verbal description, that we "come to know something that cannot be put into words" (Raffman, 1993:2). This idea of ineffability will be expanded upon in the next section, as it applies in some degree to all affective states.[56] What will be considered now is the sort of aspect of a musical utterance that can serve to elicit affective states by directly causal means.

What are examples of such directly causal aspects of musical utterances? In essence, there are two classes of characteristics that lead to affective states. The first of these has to do with those features of the musical utterance that are intrinsic to its perceived identity. The second class features those characteristics that are variable and independent of the utterance's identity. Both these classes of affect-relevant characteristics lead to affective states by directly causal means. To illustrate the distinction between these two classes, consider the musical piece 'Happy Birthday'. In any performance of 'Happy Birthday', we can change various parameters of the piece, such as the volume, the notes, the rhythm, or the tempo. However, some of these modifications change the identity of the piece. That is, when some types of changes are made, we are less inclined to recognise the piece as an instance of 'Happy Birthday'. If we change every interval from the original to some other syntactical interval (that is, the notes will all remain diatonic), a listener is unlikely to recognise the piece as 'Happy Birthday'. Indeed, if we make enough changes, it may be thought of as a new

[56] It is worth noting that the idea that music imparts knowledge of a kind that is not possible to gain through language usage is often seen as an important part of aesthetic experience. For example, Scruton (1997:200) writes that "ineffability is a mark of the aesthetic experience, in all its higher forms".

piece altogether. These changes are *identity-relevant*, and are closely related to the syntactical arrangement of the piece. Such changes also result in altered affective states.

However, there are changes that the performer can make that have no effect on the identity of the piece whatsoever. Playing 'Happy Birthday' softly or loudly, quickly or slowly, or with aggressive articulation, makes no difference to the identity of the work. It is still recognisable as 'Happy Birthday': such features are changeable without altering the identity of the work. Let us call such features *surface variables*. The distinction between identity-relevant characteristics and surface variables is best illustrated in cases where musical pieces are pre-composed, and later performed by others (such as a performance of 'Happy Birthday').[57] In such cases, the identity-relevant characteristics are those that are built into the work by the composer (form, notes, tonality, and so on), whereas the surface variables are those aspects that can be manipulated by performers (such as phrasing emphasis, articulation, dynamics, and so on). Because identity-relevant characteristics of musical utterances are closely related to musical syntax, we will now turn to a consideration of the nature of surface variables and their role in the achievement of affective states.

We have already mentioned that unexpected, loud sounds, such as the sudden orchestral chord in our Tchaikovsky example, serve to elicit affective states. Indeed, we can generalise this example to a class of phenomena on the musical surface: those of sudden changes in dynamics. More broadly, dynamics in general can be seen as an aspect of musical utterances that results in affective states. Dynamics can be used to alarm the listener, as well as gradually build or release some sort of tension that the listener perceives. Changes in dynamic level are particularly noticeable in this regard. Any person who has listened to a blues guitarist such as Stevie Ray Vaughan will be well aware of the emotional tension induced by sudden changes of volume, and any person who has listened to orchestral music will recognise the contribution of crescendo and decrescendo to the affects associated with the music. Terraced dynamics in general can serve to set the 'mood' of the music: soft music can be in turn understood as menacing, mischievous, or sad, depending on the nature of its structure and syntactical arrangement (such as tonal and melodic features).

[57] The distinction between identity-relevant characteristics and surface variables is blurred in cases of freely improvised music. However, if an improvised piece were to be repeated, it would be recognisable as a repetition provided we do not change the syntactical arrangement of the music, or alter the temporal pattern of events in such a way that it is no longer recognisable. Surface variables, such as volume, articulation, and the other variables described in the main text, can be changed without altering the identity of the repeated improvisation. (In any case, as soon as an improvisation is repeated, it is essentially a repeat of a composed work.)

A second example of a causal characteristic of musical utterances is articulation – the manner in which the music is executed. For example, a single phrase can be played with aggressive, loud accents, or it could be played in a legato fashion, or all the notes could be played staccato. All these different articulations would have different contributions to affective states. In general, short, aggressive, and loud articulations lend themselves to the elicitation of states of excitement, whereas smoother, gentler articulations are associated with calmer affects.

A third example, which depends largely on the way that a particular piece of music is put together, is the manner of phrasing. (In spoken language, phrasing's counterpart is called 'prosody'.) Musical pieces are usually put together in a way that allows for the listener to break the statement into phrases. In much tonal Western music (and tonal music in general), phrases are rounded off by cadential points, many of which demand a tonal resolution in the following phrase. The manner in which phrases are performed – where subtle (or not so subtle) emphasis is placed by the performer, or where tempo change is used as an expressive device – is a characteristic of music that acts in a directly causal manner with regard to affective mental states.[58] At this point in our discussion, the execution and emphasis of phrasing is to be distinguished from phrase structure itself. However, it will also be argued that phrase structure is also to some degree in a directly casual relationship with affective states. Phrase structure is primarily a syntactical aspect of musical utterances, and an example of a far deeper sort of affective determinant. Discussion of phrase structure will be delayed until the next chapter. For now, we are more concerned with these 'surface' phenomena that are alterable by performers, and not necessarily part of the identity or structure of a particular piece.[59]

[58] Arguably, the various techniques that constitute phrasing can be reduced to other directly causal components of musical utterances, such as emphasis via dynamic change, which notes are articulated in what manner, and tempo change. The common usage of the term 'phrasing' by musicians is the reason for its inclusion as a distinct surface variable.

[59] To illustrate this distinction, consider the case of musical works which are composed by a single person, and then are later performed by others (as opposed by improvised music). In such cases, the causal aspects of the structure that are to be seen in a deeper sense, such as the affective properties of phrase structure, are those that are built into the work by the composer (form, the notes, the key, and so on). The surface phenomena (those which we are presently discussing) are those aspects of the piece that can be manipulated by performers, such as the phrasing emphasis, the articulation, the dynamics, and so on. Note that changing the deeper affective properties of the work in performance will have a bearing on the identity of the work.

Yet another example of an aspect of musical utterances in a directly causal relation to affective states is tempo, tempo change, and the general speed at which a musical utterance is performed. In general, faster tempos convey a heightened sense of excitement, whereas slower tempos are usually more subdued. Likewise, slowing of tempo is often used as a device for relaxing tension, whereas speeding up of tempo generates excitement. This aspect of musical utterances is closely related to general rhythmic texture: in general, faster or busier rhythms, such as playing many notes in the time interval between beats, convey excitement to a greater degree than using fewer notes. Here, one must also mention the effects of syncopation as an affective device.

Timbre can also play an important part in the causality of affective states. Timbre, perceived as the quality of tone colour of an instrument,[60] is frequently used as an expressive mechanism by performers (and, of course, in specified scoring by composers). Indeed, certain timbres do have affective qualities – one simply has to think of the quality of guitar distortion used in heavy metal, or the choice of timpani and low brass to convey an effect of menace in orchestral music. These timbral choices are not arbitrary. Distortion is chosen because it sounds aggressive, low brass and timpani are often used because they sound like something menacing. These timbres result in particular affective states being elicited in the mind of the listener. Similarly, a seventh example is the experience of the phenomenal quality of pitch height. While the altering of actual pitches is effectively a change in the syntactical arrangement of the music, performers can still use the quality of timbre in combination with the pitch height to alter the phenomenal experience of pitch for the listener. High tones are often described as thin, piercing, light, or something similar; whereas low tones are often described as dark, large, cumbersome, and so on.

In summary, we find at least seven surface variables that an utterer can use to elicit an affective state in the mind of a listener. These are termed ‘surface’ variables because they are not parts of the musical utterance that the listener would consider to impinge on its identity. The utterer is able to change these aspects of the utterance without affecting parts of the music that contribute to the piece’s identity. These variables are alterable in the act of performance. The seven surface variables are:

[60] See Levitin (2006:43-45) for a basic discussion of timbre.

1. Dynamics
2. Articulation
3. Phrasing emphasis
4. Tempo
5. Rhythmic texture
6. Timbre
7. Experience of the phenomenal quality of pitch

These are not the only aspects of a musical utterance that can be used to evoke an affective state in mind of the listener. There are also aspects of the syntactical arrangement of a musical piece that can be used to cause affective states. However, it is best that further exploration of these ideas is left until we have more closely examined the syntactical aspects of music, and how they result in cognitive states. It should also be said that cognitive states can be elicited, that have affective states (and the features of the immediately represented musical utterance that directly caused those states) as their intentional objects. For example, a cognitive state can (in principle) be entertained that has some attitude toward the affective state caused by the articulation of an utterance. Such a cognitive state may be a belief that the way an utterance was phrased was indicative of a certain affect. In addition, affective states are often named – we describe a piece of music as 'sad', even though we may not consider that the same 'sadness' that is evident in another piece of music we have heard.

A bold proposition may be forwarded at this point: given the cognitive state that results from an utterance in spoken language, accompanying affective states can be triggered by the same surface means as employed by music. Loudness, articulation, phrasing (prosody), speed of delivery and rhythmic emphasis, timbre, and pitch can all be used to add to or alter the meaning of a statement in spoken language. This is a proposition already ventured by Cooke (1959), when he suggests that

> [if] we think of a group of people talking, it is obvious that, the more excited they become, the louder, quicker, and higher their voices will get; the more relaxed they become, the softer, slower, and lower they will speak… The *louder* a person speaks, the more *emphasis* he gives to what he is saying; the *quicker* he speaks, the more *animated* he is becoming; the *higher* his voice rises, the more he is *asserting* himself.
>
> (Cooke, 1959:94; his emphasis)

Cooke then invites us to consider the musical equivalents. Of course, there are exceptions to such general rules, such as creating tension (as opposed to relaxation) by using a combination of rhythmic complexity, low volume, and low register; or alternatively using loud but low booming sounds to create excitement. In these cases, Cooke suggests that in the end, for most Western music, it is the tonal treatment of phrases that decides the ultimate affect, and not secondary features of the music such as volume and rhythm (1959:95). However, as Western tonality is not universal, we may be more inclined to say that it is aspects of the syntactical arrangement of the music that have the last say with regard to the affect that is achieved. Even this is not always the case: it is doubtful whether the haunting ending of the first movement of Shostakovich's Fifth Symphony could achieve its affect if it were played *fortissimo*, even if we preserve every other aspect of its syntactical structure. These directly causal surface variables are of a very basic nature, and their pronounced inclusion triggers affective states that severely influence our perception of the overall meaning of the music.

This leaves us with an assertion that the seven surface variables listed above should be equally applicable to both musical and linguistic statements. We would expect that, if loudness and aggressive articulation led to a heightened sense of excitement in speech, the same would apply to music. This also highlights another interesting parallel between music and spoken language: the application of surface variables to musical or linguistic statements does not, to a large degree, impinge on the identity of the statement.[61] In speech, whether you are shouting or whispering makes no difference to the syntactic structure or semantic meaning of the utterance (when the words are understood literally). No surface variable will result in the listener interpreting a verb phrase as a noun phrase, or thinking that 'tree' now refers to cars. In the case of music, where structure also contributes to affect, the ultimate affective impact of surface variables would appear to be greater (such as the example of Shostakovich's Fifth Symphony quoted above). However, it is worth noting that in the case of performances of composed pieces, the performer's alteration of prescribed surface variables does not change the identity of the work. We could play a Chopin prelude with a different timbre by playing it on a church organ, but it would still be recognisable as that

[61] This is with the possible exception of the phenomenal quality of pitch in the case of intonational languages. However, the sort of phenomenal quality of pitch in spoken languages we are speaking of is the general pitch of the statement as a whole, and not the pitch relation between syllables.

particular prelude.[62] Surface variables do not impinge on syntax and semantics in language, and they do not impinge on the general identity of works in music.[63]

Before we go on to consider the case for the presence of syntax in musical statements, it would be wise to briefly expand on our previous mention of the idea of ineffability (non-propositional content) in music.

4.5 A brief discussion of ineffability and musical statements

In earlier pages, it was contended that affective states are essentially ineffable, in virtue of being non-propositional in character. Whereas cognitive states are of the sort that can be manipulated by propositionally-oriented computational processes in the mind, it was claimed that affective states cannot be subjected to such processes. The result is that such states are not available for manipulation by cognitive systems such as Language (with an upper case 'L'). This is the reason why it has been suggested that they are achieved by directly causal means. We should now further consider this charge of ineffability with regard to music.

The lengthiest philosophical discussion of the idea that music is related to the expression of ineffable states of mind is that of Raffman (1993; see also her 1988). Raffman seeks to explain the sorts of feelings that accompany the experience of music, in particular the sorts of feelings that we struggle to put into words. As was quoted above, Raffman sees the ineffability of musical experiences as the manner in which we "come to know something that cannot be put into words" (Raffman, 1993:2). That she finds it necessary to devote considerable time to the idea that music expresses things that cannot be put into words should

[62] Provided that the instrument can realise all the syntactical features of the piece, of course. For example, you cannot realise a Chopin prelude on a mbira that only plays whole-tone intervals, for instance. Furthermore, a repeat of any musical statement on the same instrument will not lose its identity as a copy of that the original statement, despite changes to surface variables. Pieces such as Cage's *Imaginary Landscape No. 4* (1951), scored for 24 randomly tuned radios, or Stockhausen's electronic music, obviously cannot be repeated on any other instruments. However, such pieces are not syntactical in the first place, as they depend on non-syntactical features (timbre) to make their artistic statements.

[63] This is not a black and white issue. A weighted aggregate of surface variables could arguably be overwhelming enough for the listener to fail to identify the piece as one that has been heard before. For example, the combination of extremely aggressive attack, short notes, high volume, and a distracting timbre, applied to 'Greensleeves', could result in the listener not identifying the piece as a performance of the folk song. The reader should also bear in mind that this is largely a perceptual, and not a philosophical, issue. The identity-relevant features of two performances of a piece are understood as those syntactical features that are perceptually identical to the listener, allowing her to conclude that the performances are of the same piece.

speak volumes for the challenge posed by this puzzle. Raffman's approach is highly influenced by computational psychology. She accepts, in principle if not in detail, the computational account of musical structure perception forwarded by Fred Lerdahl and Ray Jackendoff in the GTTM (1983). For the sake of brevity, Raffman refers to this mental schema for evaluating the structure of sonic stimuli as an 'M-grammar'. (It is to be considered as analogous to a grammar schema for language perception.) The immediate representation of a musical utterance in the mind contains elements that match up to this schema. These are elements such as the structure, the key, the intended discrete pitches, tonal relations, and so on, and are all aspects of music that we can describe propositionally. But the M-grammar schema does not account for the entire representation of the musical utterance. Raffman claims that in the total experience of music, there are parts of the representation of the musical utterance that are conscious but not verbally expressible or reportable (1988:688-689). What this implies is that there are parts of the musical representation that don't fit the M-grammar schema, but are in some manner represented in the mind. Instances of such parts of the representation include pitch and rhythm nuances that don't fit the schema exactly, but are nonetheless available to consciousness. Without the requisite musical mental machinery, such nuances cannot be represented in a way that allows for verbal expression. Hence, whatever 'falls between the cracks' of our M-grammar schema, but is available to consciousness, is what we perceive as the ineffable in musical experience.

The reader will note that Raffman's account lays similar foundations to those adopted in this study: namely, the existence of a mental capability for apprehending the structure of musical statements, the idea that there are parts of musical experience that are alterable over and above the structure of an utterance, and the idea that non-propositional content can be communicated in some manner. According to Raffman, it is these un-schematised nuances that are primarily responsible for the perceived ineffability of musical experience. However, the present study will go further by claiming that there are parts of musical structure that also contribute to perceived ineffability. Indeed, this is an answer to a concern raised by Scruton (1997:200); namely, that Raffman cannot account for the reason that most people find what is generally referred to as art music considerably more ineffably expressive than 'Three Blind Mice'.[64] If ineffability is caused only by the manner of performance, it is difficult to

[64] Of course, Scruton's project is a normative one, as he wishes to justify the claim that art music is of more aesthetic value than other sorts of music. This is exemplified in his attitude toward popular music, which for him includes not-so-popular genres such as heavy metal. Our project is not normative, but rather descriptive. I would

understand why we attribute ineffable expression to the musical piece itself. In other words, on Raffman's account, composed pieces of music are not in and of themselves expressive of ineffable content – only performances of pieces (complete with nuances) elicit ineffable experiences in the minds of listeners. Or, phrased in our own terms, Raffman is claiming that it is not the identity-relevant structural features, but rather the surface variables that result in our apprehension of the ineffable and non-propositional. This appears to be counter-intuitive.

Is it really the surface of the music, such as the fine-grained deviations from established pitch classes failing to fit our M-grammar schemas, that is responsible for our attribution of ineffability to musical experience? Lindsey (1994:576) points out that if all music were to be performed by synthesizers, hence nullifying the effect of performance nuances on the surface of the music, it is likely that a sense of ineffability will still be attributed to the pieces themselves.[65] If musical utterances conform perfectly with our M-grammar schemas with nothing 'falling between the cracks', Raffman's conclusion would have to be that experience of them will not feature any sense of ineffability whatsoever. However, it has already been contended that the structural aspects of musical utterances contribute to the formation of affective states, which are non-propositional in character and hence instances of ineffable content. This is therefore a major point of difference between this study and that of Raffman's.

Because the topic of the semantic behaviour of names is a major point of interest to analytic philosophers of language, a brief discussion of language usage with regard to non-propositional mental states is merited. Affective mental states are beyond verbal explication, being of non-propositional nature. But we are consciously aware of them, as affects (such as states of sadness, moods, or the ineffable emotions elicited by some aesthetic experiences). Raffman, as described above, attributes ineffability at base to be our awareness of the contents of perceptual experience that are too fine-grained for conceptual mental schemas. Our perceptual experience of the world in general is incredibly detailed, but many of these

---

not like to suggest that some sort of ineffable feeling fails to accompany pieces in Scruton's popular genres. In any case, Scruton sees ineffability as an aesthetic quality of great art works; I see it as a characteristic of affective states.

[65] Lindsey also points out that on Raffman's model, congenitally deaf persons are completely unable to appreciate the ineffable aspect of music in any way whatsoever. The nuances, being verbally inexpressible and only communicable by sonic means, are completely unavailable to deaf individuals. This means that no sense of ineffability could be gained from the reading of a score, as it is only surface features of the music that contribute to the sense of ineffability.

details are not taken much note of unless specific attention is directed toward them. Take, for example, our apprehension of a tree: not every leaf is consciously noted. Rather, we see a collection of leaves constituting the green part of the tree. If we wish, we can isolate individual leaves and direct our attention to them. In that way, our general perception of trees is a sort of mental shorthand, and we can call attention to detail if we like. Furthermore, it is possible to represent all of this in language.

However, can the same be said of phenomenal experience? Is it possible to describe, in language, the nature of the greenness of a particular leaf, or the smell of freshly brewed tea, or the feeling of pain? Such subjective phenomenal experiences are often referred to as *qualia* (singular, *quale*). The subjective quality of colours, tastes, tactile sensations, smells, and so on, are all qualia. In the analytic philosophy of mind, much ink has been spilt regarding these subjective experiences, particularly with regard to whether they can be explained in physical terms or not. However, what interests us is how the fundamental nature of the subjective phenomenal experience appears to be particularly difficult to frame in language. We can describe the smell of curry in physical terms, by referring to the physical constitution of the olfactory organs and the motion of particles. Some philosophers believe that, in principle, it is not possible to know what curry smells like unless you have smelt it – no amount of physical knowledge will teach you what it is like to smell curry (e.g., Jackson, 1982). Others believe that qualia are physical through and through, and therefore, if someone knows all the physical facts about smelling curry, they already know what it is like to smell it (e.g., Dennett, 1991). But can we explain the actual *smell* of curry in words?

At best, it would seem that we could say that a smell is similar to another smell, albeit not identical; and we can refer to a smell by giving it a name or referring to it demonstratively (‘*that* smell’). Is it the same with the sense of ineffability that is often attributed to musical statements? Do we apprehend ineffable ‘content’ in musical communication in a similarly subjective manner to the apprehension of qualia? It would seem that the nature of qualia and the nature of affects are similar in that they both appear to be beyond the grasp of words, being non-propositional. Hence, we cannot describe what either is like, but we are able to name the relevant mental state. We could talk of Roger’s affective state upon hearing piece *x* as X, the affective state upon hearing *y* as Y, and so on. Roger will also be able to say that X and Y are similar in that both are melancholic states, although he will not be able to say precisely in what ways they differ. Affective non-propositional states are ineffable in that

while they can be named and roughly categorised, the person experiencing the state cannot verbalise the difference between one state and another, nor render an accurate description of the state in propositional terms. In this manner, the subjective experience of the state is ineffable. We can examine a brain that is currently entertaining a non-propositional affective state and account for every neuron in propositional terms, but we cannot verbalise what it is like to experience that state ourselves.[66]

What is notable here is that while we have been saying that affective brain states are ineffable by virtue of their being non-propositional, it would be more accurate to say that it is the experience of such mental states that is ineffable (the states themselves remain propositionally expressible in terms of the physical arrangement of brain matter). Mental states are underwritten by biology, and hence, each mental state is of a physical constitution. That arrangement of brain matter is obviously describable in language after adequate empirical investigation. However, the subjective experience of being in that mental state is similar to a quale in that it cannot easily be rendered in language. It is therefore in the *experience* of affective states, that is non-propositional in nature, that the ineffability of music lies.

[66] Note that I am remaining non-committal as to the ontological status of phenomenal experiences (that is, whether qualia are physical or non-physical). What I am suggesting is that the experience is not linguistically expressible.

CHAPTER 5
MUSICAL SYNTAX AND THE ELICITATION OF MENTAL STATES

5.1 Introduction: The question of musical universals

This discussion of syntax in music is one that must, due to both constraints on knowledge and the size of the task, remain incomplete. In approaching this topic, it must be stressed that research on syntax in music not only requires a broad survey of the music of a variety of cultures, but also further understanding of how musical behaviour is realised biologically (especially neurologically). This sort of research compliments a search for musical universals. Leonard Meyer defines musical universals as "[musical] characteristics common to cultures that are unrelated ethnologically and geographically" (1960:49). To this, we should add that the concept of musical universals also encompasses those characteristics common to the experience of musical behaviour in various cultures. A theory of syntax that attempts to describe human music must take cognisance of such universals, as their validity underlies the predictive scope of any particular theory. This is a point emphasised by the study of syntax in linguistics. Ever since the advent of modern theoretical linguistics (particularly with Chomsky, 1957; 1965), a large focus of language study has been on those aspects of linguistic behaviour that are universal to all cultures. In this respect, features of syntax came under closest scrutiny, with Chomsky and his followers suggesting that the surface features of all the world's languages spring from a limited number of possible syntactic cognitive frameworks, which are innate and universal. The formalisations that have subsequently become commonplace in theoretical linguistics all owe their origins to this idea of a universal structure underlying all natural languages, even if the details of various theories have changed and are still the topic of vigorous debate (e.g., Chomsky, 1995). Thus, it is only natural that an inquiry into the nature of music *vis-à-vis* language would require a consideration of universals.

What features of musical behaviour are common to all humans, and how is it best to describe them? Do these universal features spring from cultural or biological sources? If they are biological, do they reflect constraints that are music-specific, or do they reflect the general limitations and workings of human cognitive capacities? Granted that all the cultures of the world exhibit musical behaviour, are there aspects of the manner in which music is put

together that are universal? These are important questions lying at the heart of discussions about universals, and by implication, musical syntax. In such investigations, researchers must make certain that reported universal features are carefully considered before being included in a syntactical theory. Furthermore, in any discussion of universals, it is vital that one carefully assesses what the universality of any particular trait implies. For example, the utilization of drums is a universal aspect of music-making, but one should be careful before suggesting that there is some cognitive correlate for drum-playing in particular. Preparation of food by means of cooking is also a universal feature of human societies, but few would argue that it is any sense a specifically biological behaviour. Valid universal constraints on the intelligibility of musical utterances may be simply a matter of biology, much in the same way that numerous aspects of linguistic syntactical structure reflect constraints on mental processing and the acquisition of language in infants. A comprehensive theory that addresses all universal aspects of music making in humans will have to distinguish carefully between those universal traits that result from non-cognitive biological constraints, and those that find their constitution in the way in which human brains work. A comprehensive theory will also have to make claims about which parts of the processing of music are music-specific and which are domain-general. While it would appear that headway is being made in this regard, there is still much empirical research to be conducted with regard to determining the nature of musical cognition and its status as domain-specific or -general (as evidenced by Patel, 2008a).

Some researchers have expressed the belief that the search for musical universals is futile. This is especially emphasised when the search has been conducted from an ethnomusicological perspective. Brown, Merker and Wallin write that in ethnomusicology, "universals have been the subject of great scepticism, as they are seen as smacking too much of biological determinism, and therefore of denying the importance of historical forces and cultural traditions in explaining the properties of musical systems and musical behavior [sic]" (2000:13). Ethnomusicologist George List contends that "the only universal aspect of music seems to be that most people make it. And that is about the only universal involved" (1971:402). Despite these views, scholarly opinion on musical universals has not always been negative (see, for instance, Meyer, 1960:49-50; Nettl, 2005). The idea that all human musical activity is underwritten by common principles is one that appears to have seen much currency in the first half of the 20th century, before subsequent frustration in locating those principles lead to a strong tendency to see music as culturally relative. Since the 1960s, it would appear

that the pendulum has swung both back to universals, and again in the opposite direction, to cultural relativism.[67] Currently, the predominantly cultural study of music largely shuns the idea of universal principles being applied to human musical cultures.[68] Unfortunately, this relativistic outlook sometimes seems to extend to views about music's status in cognition. The New Musicology-inspired approaches are particularly notable in this regard. For example, there has been a rejection of cognitive theories such as that of Lerdahl and Jackendoff's GTTM on grounds generally termed as postmodernist (e.g. Fink, 2001; see Cross, 2001a for a general discussion of the attitude of New Musicology toward psychological and cognitive theories concerning the experience of music). In this regard, it must be noted that a cognitive approach to musical syntax is based on the fact that the biological constitution of the members of different cultures is *not* a relative matter, as human biology is universal in all but the most negligible respects. It is on a biological basis that any approach to a common syntax of music operates, and therefore, such a theory will lay claim to universality.

It should be noted that the majority of approaches to musical universals have sought data abstracted from the products of musical behaviour – that is, from music itself. As Leonard Meyer (1960) points out, a purely descriptive view of musical universals – where different sorts of music are simply "observed, measured, classified, and compared" (p. 50) – will not provide illuminating answers about the human practice of music. Apart from the confounding variability of human music, there is also the problem of the functional reasons underlying various human behaviours. Two different cultures may have a similar scale structure, but may not share equal perception of the relative importance of particular scale degrees. Something similar can be said regarding the perception of beats in relation to periodic accents (see Cross, 2001b). This requires research to go beyond just surface structure, to a theoretical level that accounts for the role of surface phenomena in the activity of music.

Harwood (1976) notes that "[t]he *process* of understanding and engaging in musical behavior [sic] may be more universal than the *content* of musical knowledge or action" (p. 523, his emphasis). It is his suggestion that the focus on musical universals should shift from

[67] See Nettl (2005; especially pp. 42ff). It would appear that by the 1950s, relativism was in fashion. Nettl suggests that by the 1970s, the hunt for universals was back on. It can be added that since the 1990s and the advent of the New Musicology, a relativist view was established that is still dominant in much mainstream musicology and ethnomusicology.

[68] This is a particular problem when theory aspires to the level of cognitive or psychological theory. See Cross (2001a).

examining the 'music itself' to the means by which we construct and apprehend musical behaviours. The implication of this statement is that the relationship between musical universals and a theory of musical syntax should be construed as one of theoretical abstraction from (mostly) cognitive evidence. It is at this level of musical cognition that we must seek musical universals. The question of universals becomes more interesting (and potentially more fruitful) if one considers the biological capacities that allow the surface phenomena to be realised at all. Of course, if there are universals pertaining to the apprehension and production of music, it is likely that they will be somehow reflected in the musical product. However, the search for commonalities between different types of music (that is, musical pieces as the products of musical behaviour) should be guided by a bottom-up approach, from cognition to musical pieces. This is specifically where the notion of a formalised syntax comes to the fore. Universals are likely to be found at the level of cognition, in the attempt to answer questions about the mental processing of surface phenomena. Therefore, any theory of musical universals is likely to be a cognitive theory. Syntax should be evident not only in cognitive terms, but also (to a degree) in terms of the structure of individual musical performances. Of course, researchers and theorists should be wary of taking the linguistic-syntactic analogy too far. The syntax that we describe and ascribe universality to should not necessarily be considered nearly identical to that of language – an "organ of the mind", as Chomsky called the ability for language acquisition (Chomsky, 1988:131). We already know that musical cognition is a fragmentary affair, a mixture of parasitic neural networking (impinging on language, for example) and what could possibly be seen as music-specific networks (Patel, 2008a:282-298; Peretz & Coltheart, 2003). It is therefore unlikely to constitute a monolithic neural system in the brain. This filters down to the syntax we are to describe: it may be a syntax of circumstance, rather than one tailored to meet the demands of musical communication in particular. In other words, the terms in which we describe syntax are unlikely to have a direct correlation to a single monolithic module in the mind. Nonetheless, it is on a cognitive level that universals are to be identified, and from which a formalised syntactical theory will proceed.

A discussion of details of musical syntax is also likely to be one of the least 'philosophical' parts of a discussion of music's power to elicit various mental states in the minds of listeners. That is, the ultimate goal of such a project will only be achieved when a large amount of relevant empirical data has been made available, primarily of a cognitive or neurological nature. However, let us take the following to be the case for the ensuing philosophical

discussion: there are aspects of musical structure that are universal to all cultures, and that are somehow reflected in the cognitive lives of humans. Furthermore, let us assume that these features are of such a nature that we are able to formalise them as a set of syntactical principles governing the comprehensibility of musical utterances. These universal syntactical principles may spring from several sources. First, they may reflect constraints on auditory processing in general. A hypothetical universal syntactic feature may only be such in virtue of restraints on the auditory processing system in general. For example, it may be the case that the length of syntactical utterances is curtailed due to limitations on working memory, in much the same way that sentence complexity is limited by the same constraint.[69] The source of limitations may be the result of non-cognitive physical constraints as well. For example, the length of sung musical phrases is dependent on the average breath capacity of singers.[70] The means of producing musical utterances (that is, instruments) may be curtailed in what they can express because of the way that they are designed, whether in the case of a specific timbre or limitations on the actual range of sounds that an instrument can make. Another source that may result in observed universal syntactical features is a constraint imposed by the functionality of music. If we take for granted that human musicality (or some component part of it) fulfils a specific function, it may be that the character of that function directly influences the nature of its syntax. Linguistic syntax, for example, is strongly influenced by its function in conveying information unambiguously. Hence, it is organised in such a way as to maximise the effectiveness of propositional communication. It may be the case that similar constraints on musical structure may be imposed by the uses to which music is put, either now or in the distant past. Of course, such constraints do not simply limit the possibilities afforded by syntactical arrangement in music. They also serve to determine and shape that syntax in an important way, by defining the range of possible continuations of musical statements, as well as the manner in which statements can be constituted and yet remain intelligible.

Our discussion about musical syntax will span this and the next chapter. This chapter comprises a philosophical consideration of precisely how syntactical features of music contribute to the formation of cognitive and affective states in the minds of listeners. In more

[69] Indeed, this appears to be the case with the sense of tonal closure (ending a piece in the same tonal centre that it begins with). See Cook (1987) and Cross & Rohrmeier (2009) for details regarding the temporal limitations of the perception of tonal closure.

[70] The length of most instrumental melodic phrases in many types of music is roughly equivalent. I hasten to qualify this with the proviso 'melodic', as rhythm-based music independent of melody is not necessarily arranged in this manner.

colloquial terms, the present chapter will deal with how syntax (whatever that is) contributes to musical meaning. This will also include a consideration of how syntax in language compares to syntax in music, with regard to the formation of cognitive and affective mental states. The next chapter, also on syntax, is the part of the discussion that must remain incomplete. Here, we will consider several discussions of syntax in music, as well as several general observations about the future of research into musical syntax. However, the current chapter's philosophical discussion requires us to assume that there is a syntactical dimension to the phenomenon of music. Hence, we should revisit the idea of musical syntax in order to flesh it out further. Before doing so, we should pause to consider the music-theoretical notion of form, and its relation to the concept of syntactical theories of musical structure.

5.2 Form and syntax

5.2.1 *Formal and syntactical analyses of phrases*

'Syntax' and 'form' are essentially concepts utilized by theorists approaching the phenomenon of structure in music from different perspectives. 'Syntax' is a term borrowed from linguistics, whereas analysis of musical form is an activity practiced in music theory. A large part of traditional music theory and analysis is devoted to the study of form as it is manifested in musical pieces, especially with regard to large scale structure. Indeed, form has been the bread and butter of analysts for decades. Statements in form analyses can be seen as descriptions of the structure of a piece: how the piece is put together as a whole, how different chunks of music are arranged in time, and how the most important events in a musical piece are related to one another. For example, many pop songs can be said to be in verse-chorus form (alternating groups of thematic material, with each instance of the second group having a common lyric content). The jazz standard 'Autumn Leaves' is in AAB form (a repeated 'A' section, followed by a contrasting 'B' section). Modern rock and heavy metal can have extremely complex formal arrangements. However, the focus of the majority of form analyses in the academy has been the music of the Western art tradition. Here, large-scale formal arrangements range from simple binary (AB) miniatures to compound forms such as minuet and trio form and sonata form. Formal analysis also deals with smaller structures: fore-phrases, after-phrases, and motives all concern theorists that are interested in formal analysis. One of the major music-theoretical projects of the twentieth century was that

of Heinrich Schenker, who sought to explain all aspects of formal structure – from tonal structure to thematic unity in movements that were tens of minutes long – in terms of a single underlying musical idea, expanded in increasing layers of embellishment to create an entire musical piece.

Let us, for now, use 'structure' as an intermediary term. Structure is the arrangement of musical elements within a piece, whether horizontally (in time) or vertically (in space, viz. the layering of simultaneous sounds, often referred to as harmony). Thus, a single bar has a structure, as does an exposition, or an entire movement, or a chord, or three notes in succession. Up until the present moment, the term 'syntax' has been used with regard to the structure of a musical piece. To syntax was assigned several interesting features. Syntax, it was and will further be claimed, facilitates the formation of cognitive and affective states in the minds of competent listeners. Indeed, it is in virtue of the syntactical arrangement of constituent elements that a musical utterance can be said to be intelligible at all. The concept of syntax that has been proposed here thus far is a way of looking at the structure of a musical piece in cognitive terms. The music-theoretical approach to structure (that is, form analysis), does not see music from this cognitive stance. Rather, the formalisation that results from a music-theoretical perspective is more abstract in nature, focusing on the music as an entity in itself.[71] If a music-theoretical approach to structure can be said to deal with the musical intuition, it is the musical intuition of the composer: the planner of the musical piece. Form analysis examines the conscious structural choices that a composer has made. Syntax, on the other hand, approaches from the standpoint of cognition. In virtue of what is a structural arrangement of sounds considered musical by a competent listener? How do musical structures result in the perception of meaning? What is it for a musical utterance to be intelligible? Both form analysis and syntactical formalisations are approaches to theorising about the phenomenon of structure in music, and may hit upon similar explanations of musical structure. Both form and syntax are theoretical approaches that deal with the way in which musical elements are arranged to create pieces of music. It follows that each approach

[71] The reader will recall that in Ch. 3, it was mentioned that philosophical issues ('chestnuts') such as the identity of a musical work given multiple performances, as well as the ontology of a musical piece in itself, were largely reduced to problems about human perception in our theory. It is worth mentioning this again with regard to the distinctions and differences between a music-theoretic and a cognitive approach. In formal analysis, questions about the ontology of that which is studied naturally arise. However, a cognitive approach is more concerned with the production and apprehension of musical behaviour in the mind, and asking whether two performances are of a single work is likely to prompt an answer involving human memory capacities, as opposed to ontology of musical works.

will have different things to tell us about musical structure. Our primary concern has been the listener's *intuition* that music is somehow meaningful, which is an issue of human cognition. This is a feature of the listener's mental life, and we would expect a cognitive approach to tell us much about it.

To display the differences characteristic of these approaches, let us consider an example from William Rothstein's book on phrases and phrase structure in tonal music (Rothstein, 1989). Rothstein essentially argues that musical phrases are arranged in a hierarchal manner, with phrases occurring within larger phrases. We have ourselves already used the term 'phrase' in a manner suggesting that it should be considered as a perceptual unit. It has also been suggested that the perception of phrases has an 'upper limitation' in size, namely, the demands of working memory on phrase length. Furthermore, it has been suggested that phrase length is also dictated by non-musical issues, such as breath capacity. This physical limitation has a tangible effect on the manner in which phrases are performed. Hence, we would not expect our cognitive understanding of the term 'phrase' to encompass segments of music longer than that dictated by physical constraints (for example, all 32 bars of the jazz standard 'Autumn Leaves' as a single phrase, as opposed to the usual 8+8+8+8 bar interpretation of the song's phrase structure). Rothstein's approach to the concept of the phrase, on the other hand, is one of a music-theoretical persuasion. He firmly places emphasis on tonal concerns, stipulating that "a phrase should be understood as, among other things, a directed motion in time from one tonal entity to another… *If there is no tonal motion, there is no phrase*" (Rothstein, 1989:5; his emphasis).[72] Rothstein's primary concern is the Western art tradition from the eighteenth to early twentieth centuries, and therefore his interest in the harmonic features of musical phrasing is to be expected. However, off the cuff, this understanding of the phrase can lead to some surprising conclusions when considered from a cognitive standpoint. In the case of non-tonal rhythmic music, is there no phrase structure? Surely, from a perceptive standpoint, we hear rhythmic passages in electronic dance music in terms of phrases, even if the medium seems to shun tonality to a large degree.[73] Of course,

[72] From this proviso (as well Rothstein's analysis of the phrases comprising the 'Ode to Joy' melody of Beethoven's Ninth Symphony on pp. 16-17), it is likely that Rothstein will partition 'Autumn Leaves' into 8+8+8+8, according to fairly common musical intuition. However, Rothstein would also argue that there is a 'hyperphrase', 16+16 bars, and a further one, 32 bars. This is far beyond the bounds of perception of phrases as we are due to employ the term in our cognitive approach.

[73] It should be pointed out that Western art music, which is precisely what Rothstein is studying, is more concerned with creating formal structures that can be called 'narrative'. This accounts for his interest in tonal 'motion'. In contrast, many other sorts of music (such as electronic dance music and trance music) do not have this narrative, teleological concern. Rather, these sorts of music are concerned with the creation of a static state.

Rothstein's definition of the concept 'phrase' is his to stipulate, but it comes with the handicap that it can only be used to analyse tonal music in a music-theoretical manner.

The differences between a music-theoretic and a cognitive approach become more apparent if one considers Rothstein's analysis of Strauss's *Blue Danube* Waltz No. 1 (1989:5-10). In this analysis, Rothstein's contention that the determining marker of a musical segment being a phrase is tonal motion leads to the suggestion that bars 1-24, being devoid of explicit tonal movement toward a cadence, do not constitute phrases.[74] (Rothstein, in the reduction example on p. 6, goes as far as using brackets to designate phrases. Brackets are missing from the reduction of the first 24 complete bars, but are present for bars 25-32, a section with tonal motion.) This clearly deviates from what we would expect had he been utilizing a concept of the phrase somewhat similar to our own. Rothstein's analysis is at odds with the musical grouping that we *perceive* when listening to the piece. Instead of subordinating the first twenty-four bars to having less perceptive importance, we tend to hear 3 phrases of 8 bars, with the strong cadential phrase (the phrase that Rothstein brackets in his analysis as being constitutive of a 'phrase') forming a fourth 8-bar phrase. Obviously, with regard to the final phrase, harmonic figures lead us to hear these last 8 bars as a phrase. The harmonic treatment in evidence here suggests an important architectural moment in the piece (viz. a cadence, made all the more noticeable given the relative harmonic stasis of the previous three 8-bar phrases). However, the general lack of harmonic, rhythmic, and thematic action in the previous 24 bars does not prevent us from perceiving them as three separate (albeit similar) phrases, irrespective of their lack of harmonic movement. Of course, the differences between what Rothstein terms a phrase, and the perceptive concept being discussed here, are nothing more than matters of terminology. What this example shows is a tangible difference between a music-theoretic approach and a cognitive approach: with music theory, matters of perception are not as important as listener-independent structures (the 'architecture' of the piece); and in a cognitive approach, more light is thrown on how a piece is understood by a listener, or produced by a performer. At the root of the terminological dispute regarding 'phrases' is this fundamental difference in approach.

[74] Unfortunately Rothstein's reductive analysis (p. 6) and subsequent discussion is difficult to follow in the absence of bar numbers of key points in the reduction. I have therefore numbered the bars as they appear in Rothstein's examples.

Another excellent and instructive example that may help to illustrate the form-analytical and cognitive-syntactical approaches can be found in John Cage's *Sonatas and Interludes* (1948). Each of the twenty pieces that comprise this collection has a remarkably clear large-scale structure. The majority of the pieces take more than two minutes to perform, thus ensuring that there are large-scale structural features that are at a level above that of the phrase (as seen from a perceptual perspective of 'phrase'). However, what is most striking about these pieces is that the importance of pitch is drastically reduced. The majority of Western tonal music depends on pitch relations as a major contributor to intelligibility.[75] However, the impact of pitch organisation is minimal in the *Sonatas and Interludes*, as they are written for prepared piano. To 'prepare' the piano, objects such as screws, bolts, and bits of rubber are inserted between the strings, effectively turning the piano into an elaborate keyed percussion instrument. This has the result of seriously hampering the influence of pitch relations in the *Sonatas and Interludes*, as the percussive sound is issued from the piano every time a key is struck masks many tonal relations. In addition, the pieces were not written with pitch organisation as a primary concern.[76] This means that, for present purposes, we can generally consider the lack of pitch as rendering the majority of phrases in the work non-syntactical. (Of course, this example serves illustrative purposes. Thematic unity can be signalled by rhythm as well, or even timbre – as should be obvious from a repeat listening of the *Sonatas and Interludes*.)

Despite the fact that most of the phrases in the *Sonatas and Interludes* can be considered as ill-formed from a syntactical perspective (that is, taken in isolation they make little musical sense), the pieces can still be said to exhibit a very specific structural arrangement. Most of the pieces are, in fact, modelled on AB (binary) form, reminiscent of the form of most of Domenico Scarlatti's keyboard sonatas (Duckworth, 1999:85). In these binary pieces, the first and second halves are each repeated in turn, resulting in what is effectively AABB form. Consider the first Sonata as an example. In such a case, the mere fact that each section is repeated is enough for the listener to divide the piece into two halves. The first instance of the B material (bar 13) is an important structural event – important enough to be noticed by the listener and utilised as an aural landmark in the piece. This division of the piece into an A and a B section is a reflection of its large-scale structure. It is difficult to assign intelligibility to

[75] The vast majority (although not all) of the world's music depends largely on pitch relationships.
[76] The masking of pitch is not entirely perfect. In the recording that I consulted, many clear pitches could be heard. Interestingly, it would appear that Sonata IV has the note B as a tonal centre.

any of the individual sequences of sounds that make up each section. Apart from the fact that the pitch is masked due to the preparation of the piano, and cannot be depended on as a constituent of syntactical ordering, there seems to be little intrinsic relation between the sounds that make up the two sections.[77] It would appear that phrases are indiscernible, and that for all intents and purposes, the order of the sounds is non-syntactical, resulting in them sounding unmusical. Division of the sections into phrases would largely be an arbitrary affair. Hence, it would appear that we have an example of large-scale organisational structure operating independently of the intelligibility of the sound elements that make up the sections.[78] Both Cage and Scarlatti produced sonatas in binary form. However, with Cage, the individual musical utterances within each section are ill-formed; with Scarlatti, they are well-formed. It is a music-theoretical approach that will be able to tell us about the large-scale AABB structure of the piece; but from the cognitive viewpoint of syntax, the work is largely handicapped in terms of intelligibility. (That is not to say that the AABB form is not perceived as such. However, understood syntactically with regard to phrases, the work is cognitively unintelligible.)

One thing that can be noted is that in general, the longer the piece of music in question, the more likely it is that its structural arrangement is complex and hierarchical.[79] The sonatas of Cage and Scarlatti are both examples of relatively small pieces, whereas a symphonic movement by Beethoven is a much longer affair, with many discernible sections forming parts of larger sections, and so on up the hierarchy. Therefore, one would expect that a long symphonic movement by Beethoven is much more complex from a formal point of view, which is what we generally find. It is worth noting at this stage that the notion of 'phrase' that was discussed above is only hierarchical to a limited degree (if at all), but the notion of grouping is hierarchical beyond perceptive bounds. We should therefore be aware of a distinction between grouping as perceived and grouping as a higher-level abstract idea. It follows that the higher-level, abstract hierarchical notion of grouping (such as Lerdahl & Jackendoff, 1983) is best described by a form-analytical approach.

[77] Of course, it would appear that bar 1 and bar 3, being comprised of the same material, lend a degree of coherence to the opening bars of the A section.
[78] Cage's music, similarly to many other sorts of modern Western art music, is successful precisely because it plays on the listener's expectations with regard to musical syntax. The music is successful because it flouts our expectations, and herein lies the originality. Without those expectations, it would not be nearly as effective.
[79] In the view of Lerdahl & Jackendoff (1983), form and grouping structure are hierarchical through and through; and the longer the piece, the higher the hierarchy will reach.

A syntactical description of music stems from a cognitive approach to understanding musical structure. A syntactical description of musical structure brings with it certain important implications. For instance, by analogy with the study of syntax in psycholinguistics, we are looking for a description of musical intelligibility that is in some sense related to human biology.[80] A good example of how this sort of description impacts our conception of musical structure is that of physical constraints. The intelligibility of a musical phrase has an important constraint, which can be illustrated by analogy with the syntax of language. In language, syntax operates on the level of sentences. While the recursive potential of language means that a sentence can be infinitely long, there is an important fact of communication that cannot be ignored. Sentences, to be comprehendible to a listener, must occur within bounds acceptable to working memory, in terms of both self-referential complexity and the overall duration of the utterance. Therefore, to be effective, sentences have to be curtailed in terms of length and complexity. Likewise, any syntax operational in musical statements will have to be compliant with the demands of working memory to be effectual. A musical utterance could conceivably be structured recursively, theoretically allowing for an infinitely long phrase made up of chord resolutions. From the view of music theory, such a phrase is subject to formal analysis. However, a syntactical analysis would be less successful. Like the syntax of language, musical syntax is bound by the constraints of working memory. This is evidenced by experiments which show that a test subject's ability to successfully apprehend tonal closure is limited to pieces shorter than approximately a minute (Cook, 1987). Syntactical analysis would not be able to provide much insight into the structural arrangement of music beyond these bounds. As was stressed above, a syntactical analysis would be able to tell us about musical *perception*, whereas formal analysis is an approach that seeks to describe a musical piece without concern regarding the physical constraints of listeners.

It has therefore already been suggested that a syntactical description of a piece of music would be most profitable at a level relevant to the physical constraints of listeners. In language, the sentence provides a good example of the unit of syntactical analysis. What, then, is the equivalent of a 'sentence' in musical terms? What is the discrete entity within which we find syntactical rules operating? I would like to suggest that the *phrase* be considered as the locus of syntactical operations in music. Musical phrases, it will be argued,

[80] Compare this to the abstract rules used by composers to plan large-scale works. These abstract rules are not necessarily apprehended by the listener, and are not biologically determined.

are made up of *groups* that are structured by syntactical rules – to borrow a term from linguistics, the rough equivalent of a grammatical formative ("minimal syntactically functioning units"; see Chomsky, 1965:3). It is the arrangement of groups within the perceptual bounds of phrases that constitute the operation of syntax in music. We will examine more closely the notion of groups and formatives in the next chapter. For now, let us briefly consider the notion that there may be principles guiding the formation of intelligible musical utterances at a level higher than that of phrase, as seen from a cognitive standpoint.

5.2.2 *Formal and syntactical analyses of structures larger than the phrase*

We have mentioned that a token of our syntactical approach is an interest in what occurs at the level of phrases in music and sentences in language. Here, it will be proposed that there is a difference between principles governing single syntactical structures, and principles governing groups of syntactical structures. Let us take the case of syntax in language. A transformational grammar is a theory about the rules governing the intelligibility and generation of all possible *sentences* in a given language (e.g. Chomsky, 1965). However, it is not a theory about the manner in which successive sentences are strung together to form a coherent multi-sentence statement about a particular topic. When presented with more than one sentence, the same rules do not apply to the group of sentences as a whole. This paragraph is made up of many sentences. We can say that on the level of each sentence, it is syntax that governs intelligibility (what we have been referring to as well-formedness). However, that same syntax does not govern the intelligibility of the paragraph as a whole. Likewise, the syntax governing phrases in music cannot be extended to groups of phrases, such as entire sections, movements, or pieces, except in the most abstract terms. A paragraph can be made up of sentences that are perfectly syntactical in and of themselves, but that have little to do with one another, resulting in a paragraph which does not present a coherent argument. Thus, the rules that make a single musical phrase syntactical are not the same as those that apply to larger structures made up of many phrases. (Note that formal analysis provides a set of tools for understanding larger-scale structures operating beyond the level of groups. Only to a limited degree will syntactical analysis also seek to explain perception in longer pieces. It is unlikely that syntactical analysis will be able to account for conscious decisions about the finer details of structure, such as why the composer chose such-and-such a time to revisit an earlier theme, or chose to start the recapitulation section of a sonata in the

subdominant key. Consider twelve-tone pieces by Schönberg or Webern: a syntactical analysis may have little to say about the theoretical decisions that Schönberg or Webern used to create a piece, but it would have something to say about precisely in virtue of what a listener finds a particular phrase to be musically intelligible or not.)

Returning to our analogy of sentences making up a paragraph, we may be tempted say that the logical layout of a paragraph is analogous to the structure of music above the level of the phrase. The final sentence of a paragraph, for example, may be logically entailed by the previous sentences – a conclusion to an argument. There may be a particular structure to this argument. For the argument to make sense, it may be the case that sentence 3 must be related to sentences 1 and 2, and that sentence 3 must also be logically entailed by the other sentences (after Patel, 2008a:335-342). Alternatively, it may be the case that a level of semantic coherence must be achieved before the paragraph makes sense as being about a certain topic. Arguably, a similar scenario exists for musical pieces. This is, after all, where ideas of 'thematic unity' and 'coherence' come from. Whether this implication theory is the domain of syntactical or formal analysis is an interesting question. It is likely that both approaches will have something to say about this level of structural arrangement.

5.3 The robustness of the notion of 'syntax' in musical phrases

In earlier chapters, it was mentioned that we should consider music to be organised in a manner that could be considered syntactical. The argument could be summarised as follows. Musical utterances are made up of elements that are structurally discrete, and are arranged (in any given piece) in a particular order. A given piece of music is usually perceived as being meaningful in some manner. If the constituent elements of a piece of music have their order changed, the perceived meaning of the piece of music changes. Hence, ordering of elements contributes, to some qualitative degree, to the meaning of a piece of music. Furthermore, some orderings result in a severe handicap in intelligibility: they simply sound like random noise. Therefore, if ordering affects intelligibility (with some orderings being intelligible), then music is for all intents and purposes syntactical.

Of course, the idea of syntax that was used to guide the argument above was a particularly weak formulation of the concept. Under this definition, there are a multitude of phenomena

that could be considered as syntactical. The rules governing legal moves in a chess game could be considered syntactical – there are moves that are in accordance with rules that count as valid and intelligible for the game of chess. Actions in accordance with road rules could be considered syntactical in a similar manner. How is the syntax of language and that of music distinguished from the rules governing a game of chess, or a sports match? Swain (1997) suggests that the fact that music and language are combinatorial systems consisting of discrete elements marks the distinction: "The essential syntactic elements – finite discrete events bound by rules in hierarchal organizations – eliminate the pretenders to syntax. Sports and social occasions may certainly have rules, often explicit ones, but there are no basic discrete elements underlying them" (Swain, 1997:26). Swain claims three features that differentiate syntax from other rule systems. They are finite elements, hierarchical organisation, and a set of rules or principles.[81]

It is fairly obvious that linguistic syntax meets these criteria, but are they satisfied by musical utterances? It would seem that they are. Music has finite elements – notes, chords, motives, phrases – that can be combined in different ways to form utterances.[82] The arrangement of those finite elements is indeed hierarchical. From single notes, chords are formed; from notes and chords we can construct motives and phrases. And it would appear that even in the most elementary sense, those discrete elements are subject to rules. A haphazard and random configuration of notes can result in a musical statement that makes little or no sense. Hence, it would appear that according to Swain's yardstick, music is indeed syntactical.

It is worth noting a further point made by Swain. Linguistic syntax, he claims, has two broad functions: controlling the amount of information that must be processed in order to use a language effectively, and establishing relationships between words (Swain, 1997:25). With regard to information load, language would be difficult to wield without the help of syntax. Without syntax, there would be far too many words for effective communication; but with syntax, fewer words can be arranged in a manner that allows words to have multiple meanings. For example, the word 'floor' can change its meaning according to the role it is assigned in a sentence by the syntax of the language in question – say, as a verb or a noun.

[81] For those who note that chess has discrete elements (that is, individual moves) and rules, Swain correctly points out that there is no hierarchical organisation. No legal move is in and of itself more important than any other.

[82] Note that I have dispensed with a discussion of the nature of these elements themselves, as pre-syntactic entities. For a discussion of the physical manifestation of the bits that go together to form groups (that is, the sounds themselves), see Patel (2008a), Levitin (2006).

Without the help of syntax, we would require a separate word for each grammatical instance of the word 'floor' (example from Swain, 1997:27-28). Syntax also provides a means of understanding precisely what role a particular word plays in a sentence. Thus, linguistic context is important to syntax. Putting 'running' in different sentences (and not just different parts of a single sentence) changes the meaning of the word, as well as its syntactical function (*ibid.*). The same appears to be true by analogy for the various bits of sound that make up music. As we change the order in which discrete elements are combined, we are inclined to say that the meaning of the utterance has changed.

It would appear that limitation of a potentially finite catalogue of sounds takes place in musical pieces. This is analogous to the control of information intake by syntax in language. Thus, Swain's first function of syntax is present in music. But what of the second, namely the establishment of semantic relations between words? Does this have an analogue in music? According to Swain, the second role of a musical syntax is not to mediate relations between sounds, but rather to mediate tension and its resolution: "As it mediates expressed relationships in natural languages, syntax mediates the relation of tension and resolution in musical languages" (Swain, 1997:28). The notions of tension and resolution in music also found their way into the theory of Lerdahl and Jackendoff. However, I think that the tension and resolution considered by both Swain and Lerdahl and Jackendoff is the result of what is effectively the syntactical function of a particular musical event in relation to other events. One particular chord can function in different ways if placed at different locations within the phrase, and it is that placement which affects the sensation of tension and resolution.[83] Nonetheless, it would appear that in musical phrases of many genres, some events anticipate or refer to other events, in a manner analogous to the way that syntax in language mediates relationships between words. Thus, the second of Swain's functions of syntax appears to be present in musical utterances.

It has already been claimed that music is arranged *hierarchically*. The reader should note that this claim is often paired with another. It is often said that music is arranged *recursively* (e.g. Lerdahl & Jackendoff, 1983:15-16).[84] There seems to be little reason not to suggest that

[83] These sensations of tension and resolution are undoubtedly affective states. Thus, tension and resolution are the effects of causal relations.

[84] Recursion in language refers to the embedding of phrases within phrases. For example, the phrase 'who is prone to clumsiness', can be embedded in the sentence 'My cat has failed to catch any mice', resulting in 'My cat, who is prone to clumsiness, has failed to catch any mice'. Likewise, it is possible to create infinitely long

within the bounds of working memory, various aspects of music are arranged recursively. For example, a cycle of ii-V-I chords in a jazz standard can be seen as an instance of recursion (Cross & Rohrmeier, 2009). That such instances of musical recursion serve the same (or at least a similar) function to recursion in language is not entirely clear. Nonetheless, the idea that music features hierarchical structuring will also serve to distinguish its syntax from that of traffic lights and chess, and make our understanding of the term 'syntax' sufficiently robust. The claim that music is structured hierarchically implies that there are events in musical experience that are judged to be more important than others. This can be found to apply to several aspects of music. The division of a constant repetitive pulse into bars with weak and strong beats is an example of such hierarchical structuring in temporal musical experience.[85] This is especially apparent in irregular time signatures, such as 5/8 (when split into groups of say, 2+3). The fact that we hear some melodies as embellished or ornamented versions of an underlying melodic idea also shows that some structural features are perceived as hierarchically more important than others. (This last mentioned example is evidenced by the fact that when recalling melodies in experimental situations, subjects are likely to remember key events in the melodic contour that are deemed as important to the melody's identity. See Patel, 2008a:194ff.) The importance of a tonal centre in the majority of the world's music is another example of elements being heard as more important than others. This sense of tonality also accounts for the phenomenon of being able to pick out 'wrong notes' in a melody. Wrong (or sometimes referred to as 'sour') notes are pitches that are not related to those otherwise associated with the tonal centre.[86] In Western terms, we would describe wrong notes as being alien to the key that the music is being played in. The ability to hear out-of-key notes, as well as out-of-chord diatonic notes, suggests that some notes do not form part of the established hierarchy of pitches in a piece.[87] Another example is the

---

sentences by adding phrases to the beginning or end of the sentence, such as 'Mary said that her cat has failed to catch any mice', 'John asked if Mary said that her cat has failed to catch any mice', 'Peter doubted that John asked if Mary said that her cat has failed to catch any mice', and so on. Hauser, Chomsky and Fitch (2002) suggest that recursion is the single part of language cognition and computation that has no analogue in the communication systems of other animals. Cross and Rohrmeier (2009) argue that recursive structuring of chord progressions can be found in musical phrases, such as the jazz standard 'Afternoon in Paris'.

[85] Note that I am not claiming that this forms part of the syntax of music itself. However, the hierarchical arrangement of pulses into measures acts in tandem with other factors to emphasise syntactically important events in musical phrases.

[86] For a relevant perceptual study in this regard, see Trainor and Trehub (1992).

[87] Lerdahl and Jackendoff supply a good example. Imagine listening to a CD. It is likely that if a speck of dust resulted in the second chord (for example, a ii chord) of a phrase not being rendered, the listener would be less perturbed than had the speck of dust eliminated the resolution of a cadence at the end of the phrase (Lerdahl & Jackendoff, 1983:107).

automatic adjustment that listeners make when hearing out-of-tune notes. A mistuned note is heard as, for example, a flat A, and not a new pitch-class altogether.

The importance of this structural arrangement should become apparent when the reader pauses to consider if the above discussion of hierarchal ordering applies to traffic lights or the rules of most sport matches. With traffic lights, no single coloured light can be deemed as more important than any other (although ignoring a red light usually has dire consequences). Within the 'syntax' of traffic lights, one does not think of the red light as implying the green, or the previous orange light somehow influencing our interpretation of the subsequent red light. Likewise, from the stance of a football match, a goal does not rank any higher than the kickoff when assessing the well-formedness of the match.[88] With music, on the other hand, a previous event does influence the assessment of subsequent events. For example, in the jazz standard "Afternoon in Paris", bars 4 to 6 have the following chord progression (the tonic is C major): B flat min^7, E $flat^7$, A flat (example after Rohrmeier & Cross, 2009). Seen in isolation, this is a standard ii-V-I progression, much like the D min^{7b5}, G maj^7, C maj that ends the phrase (bars 8 to 9). However, the chord progression of bars 4 to 6, while seemingly resulting in the tonicisation of A flat, is *not* experienced as qualitatively similar to the progression that ends the phrase. The progression that ends the phrase is structurally more important, because it re-establishes the tonic, and it is heard as such. The cadence in A flat does not steam-roll the overall tonality of C; it must be heard in the context of C major. In hierarchical terms, it is subservient to the final cadence. This is a strong perception, and we expect it to be reflected in any formalisation of musical syntax as it would appear to directly affect our apprehension of the meaning of a musical utterance.

It is clear that music boasts some sort of syntax in virtue of the fact that the ordering of constituent elements has an impact on what listeners consider to be the meaning of an utterance. This ordering is hierarchical, with some events being of greater structural significance than others. The grammatical structuring of language is also hierarchical, although it is not yet entirely clear whether this is the same sort of hierarchy as we find in musical structure. (That is, it is hard to say that there are elements in a linguistic utterance that are structurally more important than others – although phrase heads do come to mind.

[88] Of course, in language and music, the finite elements themselves have the potential to be combined in a syntactical manner (in the way the 'ran' is usually used as a verb, and not a noun; or the tension created by the leading note in much Western music). This 'vocabulary' is not a feature of traffic lights or chess moves.

Comparison of the hierarchy between the syntactical dimensions of music and language is blurred by the semantic dependency of linguistic grammar.) Both music and language consist of discrete elements that are ordered in some way. There are orderings of constituent elements that appear to make no sense at all, and that are perceived as musically unsound – think of, for example, a toddler bashing randomly on the keys of a piano. It would thus appear that some sort of set of rules or principles guides the formation of intelligible musical utterances. In a similar manner, the ordering of constituent elements in language (formatives and words) has an impact on the meaning of utterances over and above the meaning of those individual elements in isolation. And, of course, there are orderings of words that make absolutely no sense at all. Thus, the rule-bound combination of discrete elements in a hierarchical fashion is common to both language and music. Both systems are organised syntactically.

It is far from established precisely how closely the syntax of language and music are related. However, it seems safe to say that in a broad sense, both language and music are governed by syntactical principles that differentiate them as communicative systems from traffic lights and flag signals. This can act as a premise in the following discussion of how syntax affects the formation of cognitive and affective states in music and language.

(The reader will note that the conditions laid down for syntax in the discussion above superficially apply to what is commonly designated as large-scale form. Formal arrangement is hierarchical: there appear to be parts of form that stand out as architecturally more significant than others. In sonata form, we may say that an introduction or a linking passage between principal themes is subservient to those themes.[89] Large formal structures are arguably made up of discrete units: we are able to distinguish A sections and B sections, recapitulations and development sections, and so on. However, it is not altogether clear that in all cases these can be recombined as discrete units according to a set of rules. In a sonata by Beethoven, it would be inconceivable to substitute the recapitulation of one sonata with that of another in the same key. Thematic coherence between the sections would make this a futile exercise that would distort any syntax- or form-related meaning. However, the case becomes more plausible if we consider substituting a trio section into an 18th century minuet

[89] We may be able to distinguish between themes and linking passages by the 'density of meaning'. Rephrased in terms such as those we have been using, we may say that a large part of our ability to distinguish principal themes from linking and introductory passages is the ability of such material to evoke cognitive and affective states in the mind of the listener to a more significant degree than we find with linking passages.

and trio, preserving the correct key relationships. With regard to rules, it could be argued that the units that make up large formal structures are bound by some sort of rule system, possibly involving thematic coherence. We will return to this problem in the next chapter. As the reader may have guessed, the distinction that renders a syntactical approach distinct from one of formal analysis is the constraint on the perception of large-scale structural relationships. It will become apparent that other aspects of musical syntax, such as metric-pattern regularity, do not impose themselves at the level of large-scale structure.)

5.4 Syntax and the formation of cognitive and affective states

In earlier chapters, it was suggested that the apprehension of both music and spoken language results in the formation of cognitive and affective states in the minds of competent listeners. The co-existence of a multitude of such states, elicited by virtue of the musical stimulus and conscious consideration thereof, give rise to the perception that music is somehow meaningful. When combined with semantics, the elicitation of such states is characteristic of linguistic communication via speech. In addition to a lack of semantics, it has also been suggested several times that what distinguishes music from language is the fact that it tends to elicit affective states to a far greater degree. Discussion of how music can elicit cognitive states was delayed, as was a discussion of how it is possible that structural elements of music might give rise to affective states. The reader will recall that in language, structure is mostly dictated by the demands of syntax, and that syntax is a means to semantic ends. Structure in language is subservient to efficacy of communication. The structure of everyday speech has little to do with eliciting affective states. The primary purpose of the organization of words in an utterance is to convey semantic meaning reliably, and thus allow for the formation of cognitive states. We do not use linguistic structure to induce affective states – far easier would be the use of articulation, prosody, and other factors similar to those performative surface variables outlined in the previous chapter. Of course, in poetry and prose, writers often utilize linguistic structure for aesthetic means. Any person who has read or seen a play by Shakespeare will be able to attest to the fact that the manner in which language is wielded is an important conveyer of affect. Had Shakespeare preserved the semantic meaning of many of his sentences, but altered their grammatical construction, the artistic result would not have been the same.

I would like to contend that this part of the linguistic analogy does not extend to music. That is, in music, structure *does* in fact lead to the formation of affective states on a regular basis. Indeed, it is the presence of these affective states that have led to music often being described as an 'aesthetic' medium, as opposed to the largely communicative uses to which language is put. With everyday linguistic statements, one would not expect the manner in which words are combined to play a pivotal role in arousing affective states. However, with music, it seems to be the case that various structural arrangements do indeed contribute to the elicitation of affective states. It is difficult to imagine how a particular musical statement could achieve the same affective state if it has been changed in some structural way.

To return once again to our example of Tchaikovsky's loud orchestral chord, it is hard to see how the composer could have made the music as affective without that specific musical figure. A soft but dissonant chord would simply not be a satisfactory substitute: it would not achieve the same ends. Had the loud chord been placed elsewhere (say, halfway through the next phrase), the resultant affective state in the mind of the listener would not be the same. Compare this scenario to the semantic content of a sentence informing you of something. The semantic content of a linguistic utterance can result in an affective state, by triggering a cognitive state that then results in an affect. The listener apprehends the actual semantic meaning that constitutes the cognitive state, and then realises the implication of that meaning. This realisation can trigger an affective state (e.g., of alarm or sadness). Let us say that John has just received a telephone call. The lady on the other end of the line stated: "Congratulations, you have just won the lottery draw". Now, the resultant affective state was not achieved by directly causal means. Rather, in John's mind, a cognitive state is constituted by the apprehension of a proposition. This cognitive state relating to winning the lottery may be something like John's belief that he has just won the lottery. Now, John is an excitable fellow, and having achieved the belief that he has won the lottery, is subject to an affective state that we can call 'delight'. Where does this affective state of delight come from? It is not the result of any performative aspect of the caller's voice. Rather, it results from the cognitive state of believing that he has won the lottery. Furthermore, that affective state in the mind of John would not be different had the structure of the sentence changed: "The lottery draw has been won by you, congratulations". Thus, it is the semantics and not the structure of the sentence that has, via a chain from cognitive to affective states, resulted in the affect of

delight.[90] Now consider something analogous in the case of music. If we had changed the order (that is, the structure) of the notes in a piece of music, such as placing Tchaikovsky's loud chord at the end of the phrase instead of the beginning, or swapping the V and I chords at the end of a Bach chorale, we would not arrive at the same affective state had we not made a change. Reflection on such an elementary example shows how important structure is to affect in music. In everyday language, affective states arise regularly from semantics, but only in exceptional circumstances from the structure of the sentence involved. In music, structure is vital to affect. The basis on which a piece of music is constructed is important and holds consequences for the resultant affective states in the mind of the listener.

Above, it was mentioned that with language, semantics can lead to affective states by virtue of a chain starting with a cognitive state. We will examine such a chain below, and consider the difference in the nature of the causality of affective states that are the result of chains of other states, and those that are caused by properties of the sounds themselves as they are immediately represented in the mind, in a manner similar to automatic reaction. Thereafter, the discussion will turn to the structure of music and its ability to elicit affective states. At this point, I would once again like to emphasise that I am not able to offer a complete account of syntax in music. To do so is not only beyond the scope of this study, but is also a daunting task for any musicologist. What would be required would be an immense amount of ethnomusicological and ethnographic evidence, tied to data from cognitive musicology in a theoretically meaningful manner. When one considers that debates about the fundamental nature and theoretical description of linguistic grammar are still lively some five decades since modern linguistics took off, the task for musicology seems to be imposing. However, it is a task that evolutionary musicology in particular will have to face up to in the future, and I think that it is not beyond reasonable expectation that fundamental theories can be forwarded already. Nonetheless, we will step off from the premise stated above, namely, that music can indeed be said to have a syntactical dimension.

[90] Let us ignore, at present, the deep structure of these two sentences, as posited by theories in transformational grammar. In such theories, the two sentences would indeed be seen as identical in deep structure. The fact that these two sentences share a deep structure explains why they mean the same thing.

5.5 Mechanisms of causality

5.5.1 *Directly causal mechanisms and cognitive-causal links in language*

We will now consider the claim that cognitive states can themselves induce affective states. This is a subtle point that is best explained in terms of the processes of causality under which affective states are produced in the minds of apprehenders of musical and linguistic utterances. For ease of explanation, we should once again consider the case of language. The semantic content of a sentence is determined by the meaning of the words in a particular arrangement. In addition to the lottery example described in the previous section, consider the following hypothetical case. A robber bursts into a bank, shouting profanities and making demands at a high volume. In addition to the robber informing the teller (in virtue of the semantic meaning of the words that he uses) that he is robbing the bank, his aggressive tone serves to elicit an affective state in the mind of the teller. In this situation, the aggressive tone, as represented in the mind of the teller, is in a directly causal relationship with the affective state. If the robber had burst in and shouted aggressively in a foreign language, there still would have been an affect experienced by the bank teller, such as an affect of fear or alarm, resulting from the manner in which the utterance was delivered.[91]

Let us now try to imagine this example with an absolute minimum of utterance parameters that act in a directly causal manner. Imagine that the bank robber stood in the queue, and when he got to the teller, he slipped a note under the glass window that read as follows: "This is a bank robbery. If you raise an alarm, your life will be in danger, because I have a gun. Hand over the cash in silence." Few would deny that our hypothetical bank teller will experience an affective state of fear. But without a directly causal dimension to the utterance, how is the teller's affective state of fear achieved? After all, there was no tone of voice or aggression conveyed on the written note.[92] The answer has already been put forward in the previous section: it is the apprehension of the semantic meaning of the utterance (that is, a cognitive state), that triggers the corresponding affective state.

[91] It should be added that even if the bank teller does not put two and two together and realise that a robbery is in progress, the inflection of the robber's voice will still elicit an affective state of some kind. This is a prediction made by this theory: even if semantic and pragmatic concerns play no part in the apprehension of an utterance, the manner of delivery of the utterance can and will result in an affective state by directly causal means.

[92] Of course, I am not denying that affective states can be elicited by the features of visual-based communication. Think of pictures and paintings, for example.

It is clear, then, that the semantic meaning of a sentence can result in the formation of an affective state. For an affective state to be realised in the mind of the apprehender of an utterance, a causal relationship is at some point required. At some point in the chain of cognitive states (beliefs and so forth), there needs to be a causal link that results in the affective state. However, because this causal link is in a chain of cognitive states, we can say that it is not *directly causal*. This differentiates such an affective state's causal history from that of affective states elicited by (for example) the inflection of the voice. In the case of the first bank robber example, where the robber shouts aggressively, the affective state in the mind of the teller is the result of a directly causal process. However, in the example with the note, and all other examples that depend on semantics to elicit the affective state, let us say that a *cognitive-causal link* is in operation. Hence, there are two ways in which an affective state can be achieved: (a) in a *directly causal* manner, with no propositional or deductive mediation between the stimulus (as it is immediately represented in the mind) and the resultant affective state. The affect is involuntary, similar to any other hard-wired reaction (given that the listener is competent and is not conditioned to react differently to this stimulus); and (b) the affective state is achieved by a *cognitive-causal* link, where a cognitive state results in the elicitation of an affective state, as in the case of semantics causing affects.

5.5.2 *The elicitation of cognitive states from musical utterances*

We saw in the last chapter that a primary mental manifestation of linguistic utterances was the cognitive state, often by means of semantics. Semantics also leads to the elicitation of affective states by a causal path that we have termed a cognitive-causal link. In addition, in cases where characteristics of performance such as voice inflection and prosody affect meaning, we noted that spoken language is capable of resulting in directly causal affective states. We must now turn to the question of whether music can elicit cognitive states. The claim has already been made that music does have this power to some degree. What remains is for us to reconsider the state of the argument with regard to directly causal affective states, and affective states that are the result of cognitive-causal links. In other words, we must consider whether music is able to elicit affective states without the benefit of intermediary cognitive states. This is where music's lack of semantics becomes most apparent. The reader may recall that we have already ruled out the idea that music has a semantic dimension, as

'semantics' defined as per our modified Tarskian definition (SDS) applies only in cases of arbitrary symbolism. In language, the cognitive-causal link was the result of semantics: namely, the self-reflexive apprehension of the cognitive state itself leads to an affective state. One of language's peculiar powers is the use of otherwise arbitrary symbols to convey messages that can have a major bearing on the psychological state of the listener. But is an analogous scenario applicable to music? Without semantics, can the sorts of cognitive states that are the result of musical experience lead to affective states via a cognitive-causal link? For this to be the case, we are looking for a scenario where a chain of cognitive states elicited by musical experience can result in an affective state. Can such a scenario be found?

One should note that, as pointed out previously, semantics does not do all the work with regard to the cognitive states resulting from language use. For example, in the apprehension of any utterance with an affective component, a resultant cognitive state could be the apprehender's awareness that she is experiencing an affect. (The verbal report would be of the form, "I am experiencing an affective state". Note that an accurate and complete description of that affective state itself is not possible, as affective states are not expressible in linguistic terms. The affective state can, however, be given a name, such as 'fear' or 'alarm' or 'foreboding', and henceforth be used propositionally.) Other cognitive states could be beliefs about the psychological state of the utterer, the state of fearing whatever is implied by the semantic meaning of the sentence,[93] or cognitive states resulting from the actual phenomenal experience of the statement ("The inflection of the utterance was aggressive", as opposed to the affective state induced by the property of aggressive inflection).[94]

With regard to the phenomenal experience of an utterance, it is important to take note of the distinction between cognitive states that are effable in principle, and those that are effable *de facto.* Let us suppose that encoded in the mind is a mechanism that plots the inflection of an utterance in physical terms – that is, in terms of the frequency of the fundamental and the partials of each note in the utterance, a finely grained and quantified representation of the temporal and rhythmic character of the utterance, a highly detailed representation of amplitude, and so on. In such a case, the apprehender may not be consciously privy to all the details of an otherwise automatic mental analytical operation. Rather, the apprehender may

[93] Remember that 'fear' is an affective state, but the propositional attitude 'X fears that *p*' is a cognitive state.
[94] Note, too, that in the case of cognitive states resulting from the phenomenal experience of the utterance, some states may be conscious and others unconscious. This is discussed in the main text (the distinction between effability in principle and effability *de facto*).

simply have conscious access to a phenomenal judgement that is the end product of the mental analysis. The apprehender may not be able to say "The amplitude of F1 at time *t* was *x*", but he may be able to say "The middle of the sentence was loud". This phenomenal judgement is a sort of mental shorthand for the complex quantified representation that is apprehended. Nonetheless, the quantified representation that is unconscious is *in principle* an effable (and therefore cognitive) state, as it can be (in principle) represented in language. Such representations are, after all, the very stuff of propositional-computational complexes.

Musical utterances can result in cognitive states in a manner analogous to that of language, as described above. Along with the elicitation of affective states, the apprehender of the musical utterance can form a cognitive state about a particular affective state. While we may not be able to describe in language a particular affect associated with a piece of music, we are able to entertain a cognitive state that has the affect as its object. The apprehender will be able to report on a belief about the affect in question, for example. The apprehender is able to utter general statements about the affect (e.g., "This music makes me feel sad"), without being able to ultimately describe the affect in question (the state of sadness).

The idea of states that are effable *de facto* and effable in principle apply to the case of music as well. For instance, the mind may represent a tone quite literally, as the fundamental frequency and several partials in a detailed manner. However, the representation in such a fine-grained format is probably not available to consciousness – a contention I share with Raffman (1988; 1993). Instead, we perceive the sensations we call 'pitch' and 'timbre'. Nonetheless, the frequencies and relative volume of the fundamental and partials are, in principle, effable. Given an apparatus that can match the detail of the format used in this subconscious representation, all available details will be communicable in language. However, what bubbles up to consciousness is not a fine-grained representation of the harmonic spectrum of a note, but rather a perception of pitch and timbre. This is precisely the same case as above: namely, mental shorthand for a quantified representation in the mind of the listener. However, a statement such as "The first E-flat in this piece was played slightly sharp" results from a cognitive state, and is of course perfectly effable *de facto*, and the fine-grained quantified representation in the mind is effable in principle. Indeed, statements about the surface details of a musical piece are possible. The listener can comment on form, key, tuning, instrumentation, and so on, as these are aspects of the musical experience that are manifested in the mind as cognitive states. There seems to be little reason to doubt that it is

cognitive states that are at the root of statements about surface details and the relationship of the constituent elements of a musical piece (such as form). Our beliefs about the structural surface of a piece of music are reportable, even when we may have overlooked some aspect of form, or have failed to see some formal relationship that the composer intended us to see.

It is in this manner that music results in cognitive states. But can these cognitive states result in affective states in the same way that semantics in language can lead to affective states? Can we experience a music-induced cognitive state that leads to an affective state? I would like to argue that cognitive-causal chains are a unique characteristic of semantics, and are therefore one would *not* expect to find them in music. In other words, I am contending that in music, a chain of cognitive states relating to the music itself *cannot* result in an affective state. Two immediate problems face this claim. First, we have noted that the syntactical properties of a musical utterance, which can themselves be represented in the mind as cognitive states, result in the formation of affective states. This appears to be in contradiction with the above contention that musical experience does not feature cognitive-causal chains. Second, the seasoned listener of music will readily tell us that music often does refer to things beyond itself; hence, in some sense, music has reference and therefore a semantic dimension. This is even true in music that has no programme. Parody, quotation, and arguably genre-specific musical figures are all cases of music referring beyond itself, to situations beyond the bounds of the immediate musical experience. While these two objections are easy to formulate separately, they are both answered with reference to music's lack of semantics. Let us consider these challenges in turn.

It has been argued that the syntactical arrangement of the constituent elements of a musical utterance have an impact on meaning. It has also been argued that syntactical arrangements have an impact on both cognitive and affective states. A particular chord has a structure which can be explained in terms of syntax, and is perfectly expressible in language. One can describe the notes that make up the chord, their individual fundamental frequencies, the quality of the chord in music-theoretical terms, and so on. These are all represented in the mind as cognitive states. But that chord, placed in its syntactic context, may also have a tangible effect on the formation of affective states. The composer, in using that specific chord at that specific time, is able to induce an affect in the listener. So how can it be that features of the musical utterance that are represented in the mind as cognitive states do not result in affects? The answer, I think, has to do with a distinction between the *experience* of music and

propositional knowledge about music. I would like to contend that there is no link between these cognitive states and the affective states resulting from a particular musical figure. Rather, the affective state is not the result of cognitive apprehension of the musical figure in question – that is, the affect is not caused by anything propositionally known about the musical utterance. The affect is instead the result of a directly causal relationship between the experience of the musical figure (i.e. the immediate representation of the stimulus), and the affective state. This contention amounts to the claim that a listener can be privy to all available propositional knowledge with regard to a musical utterance, but will only be able to form an affective state if the utterance is actively *experienced*. That is, a causal relationship can only hold if the listener has experienced a musical stimulus. A single musical figure can be explained in propositional terms, but without the benefit of experiencing the musical figure itself, no causal relationship can result.[95] It is not the cognitive realization of the structure that leads to the affective state, but rather the experience of the structure. This is a directly causal relationship between the immediate representation of the stimulus in the listener's mind and the resultant mental state.

Another way to consider this scenario is by comparison with language. With semantic states (which are cognitive states), deduction and induction is possible. It is possible to form cognitive states with regard to linguistic utterances, and from the semantics of the utterance, infer other cognitive states. It is this scenario that resulted in the postulation of cognitive-causal links: the realisation of the implications of the semantic meaning of an utterance can ultimately result in an affect. The reading of an otherwise causally neutral written note can result in the realisation that one is late for a meeting, or has forgotten a birthday, or is about to undergo a silent bank robbery. This inference from the original semantic meaning of the utterance to a cognitive state resulting in affect constitutes a cognitive-causal link. Now consider music. Except under exceptional circumstances, music has no equivalent to semantics. Mental states resulting from musical experience cannot be used to infer other mental states. Having apprehended the structure of a piece of music, one cannot use this information to infer other cognitive states, which will then result in the realisation of an affective state. Induction and deduction are not possible without semantic content, and it is

[95] It may help the reader to recall that it is the affective mental state that is inexpressible in language, not the musical figure itself. Hence, aspects of the musical figure can be represented in a cognitive state, but the affect that results from experiencing the musical figure is not effable.

contended here that semantic content is precisely the sort of thing that music does not have.[96] Cognitive-causal chains, with cognitive states leading to affects, are what constitute deduction and induction in the mind. If musical experience does not result in semantic states, it cannot result in deduction and induction, and it cannot feature cognitive-causal chains of mental states. The sorts of cognitive states that are present in musical experience are not semantic. They are at best cognitive states *about the music itself*, and not about anything beyond the music. The only time that this is not the case is when musical figures are agreed by public convention to refer to something beyond the music. In such cases, music is being attributed referential properties, and is for all intents and purposes acting in a semantic manner akin to language. This leads us to the second objection.

The second objection to the idea that musical experience does not feature cognitive-causal links was that musical pieces often do feature figures that refer to things beyond music. A good example is Tchaikovsky's *1812 Overture*, which features quotations of the Russian and French national anthems, as a direct reference to the French and Russian armies involved in Napoleon's disastrous retreat from Russia in 1812. Here, musical figures unambiguously refer to things beyond music. In a hypothetical situation, a particularly misinformed and paranoid Russian patriot, oblivious to the Overture's program, may form a cognitive state regarding the two anthems being used, and interpret the music as somehow signalling conflict between the two nations. Henceforth, he may arrive at an affective state of 'fear' when he remembers that his neighbour is French. This appears to a case of a musical experience where a cognitive state begets an affect. Therefore, it would appear that a cognitive-causal chain is indeed in operation, and that a music-induced cognitive state has resulted in an affective state. The answer to this objection once again lies in the fact that music does not boast semantics. In purely absolute music, the notes and structures of musical utterances would not refer to

[96] There is, of course, the contention that if one syntactic event implies another whose experience will result in an affective state, a cognitive-causal chain is present (for example, a particular chord implies a particular resolution, that when experienced, results in an affective state). However, I believe that this argument can be countered in three ways. The first counter-argument is admittedly rather weak: in such cases, the former syntactic event, which implies the latter resulting in an affective state, is acting in a manner analogous to semantics (i.e. a semantic scenario is actually present). That is, it is referring to the latter syntactic event, and by these semantic means, is resulting in a cognitive-causal link. The second counter-argument is that the rules and principles governing the implication of one chord by another may be so weak that several resolutions are possible, and therefore, no single syntactic event is implied by another. There are a multiplicity of possible implications that a single syntactic event may imply; therefore, it cannot be said to be referring to any future syntactic event in particular. The third counter-argument is related to the second, in that it may be that only the experience of the latter syntactic event will result in an affective state. In other words, without the experience of one of the possible implied syntactic events, no affect is experienced. I believe this case to be the strongest argument against the contention, and a similar argument is featured in the main text.

anything. However, exceptional and artificial circumstances can be thought out with non-absolute music. These are circumstances of public convention, and are equivalent to the public conventions that result in the words of a language having particular referents (as manifested in sounds and in written form). Public convention allows us to say that 'cat' refers to furry four-footed felines, and that a certain combination of sound structures, by virtue of being an anthem, refers to either France or Russia. In other words, when a musical figure is agreed by convention to refer to something beyond music, it is not simply operating in a manner analogous to semantics in language. It *is* semantics. There is no qualitative difference here: when musical figures have conventional, non-musical meanings, music is being treated like language. In these cases, music has been attributed a semantic dimension by the listener by virtue of public convention. And where there is semantics, there are cognitive-causal links. As long a musical utterance is semantics-free, there will be no cognitive-causal chains possible, and cognitive states will not lead to affective states.

Of course, we often attribute conventional, public meanings to musical figures. In this regard, music is given semantic properties and functions like language. In practice, this is a regular occurrence. It is on these grounds that programmatic music operates. Quotation and parody of musical style also depend on musical figures being treated semantically. However, if we were able to imagine an idealised case without such semantic references, we would find that there are no cognitive-casual links in musical experience. It is only when music is used as language that it boasts this linguistic property at all. Otherwise, the absence of cognitive-causal links is a major disanalogy between music and language as communicative media.

CHAPTER 6
SUGGESTIONS REGARDING THE FUNDAMENTAL NATURE OF MUSICAL SYNTAX

6.1 Introduction

> [Similarities between music and language] include the existence of multiple levels of combinatorial organization, hierarchical (and recursive) structuring between elements in sequences, grammatical categories that can be filled by different physical entities, relationships of structure versus elaboration, and context-dependent grammatical functions involving interdependent relations between elements. These similarities are interesting because they suggest basic principles of syntactic organization employed by the human mind.
>
> (Patel, 2008a:267)

In the previous chapter, it was argued that from a philosophical point of view it is profitable to think of music as being organised syntactically. Changes in the ordering of the constituents of a musical utterance result in changes in the meaning perceived by the listener. Hence, it would appear that the principles governing the structure of a musical statement contribute to its meaning, and *ipso facto*, that music boasts some sort of syntactic structuring. It was also stressed in the previous chapter that arriving at a definitive description of musical syntax is a task beyond the means of the current study (and, arguably, the current state of empirical knowledge, especially with regard to our present neuroscientific understanding of musical cognition). The purpose of this chapter is to discuss the state of research surrounding musical syntax, as well as the nature of the enterprise of formalising relevant empirical data to best represent musical syntax theoretically. As has already been suggested, the most comprehensive and influential formalisation of a musical syntax is Lerdahl and Jackendoff's *A Generative Theory of Tonal Music* (1983; hereafter referred to as 'GTTM'). Although the theory has occasionally come under fire from musicologists inclined toward the 'New Musicology' paradigm (e.g. Fink, 2001; Cook, 2001b), its impact on studies of music cognition has been large and researchers in the field ignore it at their peril.[97]

[97] Approaching from a postmodernist perspective, Fink claims that the 'surface-depth' approaches he associates with the GTTM '[seem] arrogant and naïve, if not simply hopeless' (2001:103). In the same volume, Nicholas Cook suggests that the structuralist approach he sees inherent in the GTTM has the effect of '[explaining] music

The GTTM is not only important with respect to the details of the theoretical formalisation it proposes. It also represents an important disciplinary milestone. Lerdahl and Jackendoff were inspired to a theory of this kind due to the interest generated by Leonard Bernstein's lectures at Harvard in 1973 (published as Bernstein, 1976). Bernstein's lectures were in turn inspired by linguist Noam Chomsky's transformational generative grammar theories of the 1950s and 1960s (especially Chomsky, 1957; 1965). The GTTM is a complex theory that seeks to describe the unconscious abstraction that is made in the mind of an experienced listener when encountering a piece of tonal music. Lerdahl and Jackendoff use the term 'generative' to imply that the application of a set number of rules could result in a listener arriving at a structural abstraction of any given tonal piece; or, as Patel (2008a:240) puts it, "the use of formal procedures to generate a structural description of a given musical piece". However, it is important to realise that while the ultimate inspiration for the GTTM was the work of Chomsky, Lerdahl and Jackendoff sound a stern warning for those embarking on comparative studies of music and language: "pointing out superficial analogies between music and language, with or without the help of generative grammar, is an old and largely futile game" (Lerdahl & Jackendoff, 1983:5). The GTTM does not simply attempt to apply the tools and techniques of generative grammar theories to music cognition. There are none of the strained comparisons between the elements of music and parts of speech, such as nouns and verbs, which can be found in Bernstein's lectures. The famous syntactic trees that are featured in the GTTM do not relate to music in the manner which Chomsky and other linguists mapped the syntactical constituents of sentences, but rather illustrate the hierarchy of events in terms of tension and resolution and structural importance (Patel, 2008a:241, 263; Lerdahl, 2009).

What sorts of similarities exist between music and language that led to Lerdahl and Jackendoff being inspired by Chomskian linguistics? First, both music and language are arranged according to hierarchies, with structure being realised on multiple levels. In a similar manner to which language forms sentences from phrases, phrases from words, and words from morphemes; musical utterances are composed of phrases, chords, and individual notes, in hierarchical fashion. Lerdahl and Jackendoff also contend that music features recursion (Lerdahl & Jackendoff, 1983:14-16). This is a contention that is shared by several

---

without musicians' (Cook, 2001b:242). Cook also claims that the commentary on the intelligibility of modern musical composition in Lerdahl (1988), which is based on Lerdahl and Jackendoff (1983), 'slips imperceptibly from description to prescription, so reinforcing the hegemony of theory' (Cook, 2001b:252).

authors on the topic of music cognition (Cross & Rohrmeier, 2009; Jackendoff, 2009; Mithen, 2005:17; Cross, 2005). The idea that structure governs the intelligibility of musical statements, and that structure can be elaborated hierarchically, is also an important similarity that is brought out by the GTTM (as is pointed out by Patel, 2008a:267). This implies that a musical statement is heard in structural terms by the listener, and all the surface features of that statement are heard in relation to key structural features.[98] Thus, it would appear that the syntax of music is complex enough to merit a comparison with that of language.

In a seminal article on the conceptual issues surrounding the investigation into the evolutionary origins of language, Hauser, Chomsky and Fitch (2002) note that the investigation of those aspects of our linguistic capacities that appear to be human-specific have to do with our syntactic capabilities.[99] Following suit, Patel (2008a:267) has noted that the presence of another syntactic system in the mind (that is, music) would demand attention from cognitive scientists. This idea is present in the GTTM, with its cognitive approach that attempts replicate the methodology that has been so successful in the study of language. However, as has already been noted, the GTTM was not simply an exercise in the application of theory from generative linguistics to musical perception. There are major differences between music and language that are reflected in the formalisation, differences such as the lack of semantics characteristic of music and the resulting lack of grammatical categories. The true influence of the Chomskian influence is most clearly recognised as a matter of methodological style. In the words of Lerdahl, reflecting on the 25th anniversary of the GTTM,

> it was Chomsky's way of framing issues that attracted us: the supposition of specialized mental capacities, the belief that they could be studied vigorously by investigating the structure of their outputs, the distinction between an idealized capacity and its external and often accidental manifestations, the idea of a limited set of principles or rules that could generate a potentially infinite set of outputs, and the possibility that some of these principles might be unvarying beneath a capacity's many cultural manifestations.
>
> (Lerdahl, 2009:187)

[98] The reader may profitably draw an analogy with the idea of deep structure in linguistics.
[99] Although, Hauser, Chomsky, and Fitch hypothesize that even these syntactical abilities may be the result of aspects of cognition that are employed in different ways in other animals.

This undoubtedly places the GTTM within the class of theories that was described as 'cognitive' in the previous chapter (as opposed to 'music-theoretical').[100] It is this approach that is characteristic of theories focussing on music perception. The GTTM's perceptual question is as follows: what are the musical intuitions of a listener who is experienced in a genre of music when he or she listens to a musical piece? It is worth noting that in the GTTM, no assumption is made as to whether there are aspects of musical cognition that are music-specific innate mechanisms, or if innate learning mechanisms (whether music-specific or domain-general) are present in the mind of the listener. Rather, the theory accounts for experienced listeners, who have had adequate exposure to the tonal idiom as to allow them to 'understand' the music. At first blush, this proviso and the fact that the music handled is tonal, may lead the reader to believe that the GTTM has relatively narrow concerns (that is, the cognition of a specific target group within a specific musical genre). Nonetheless, Lerdahl and Jackendoff propose that many of the rules they postulate are universal, and can be found in all the musical cultures of the world. Some of these rules do indeed appear in radically different types of music. For instance, Temperley (2000) has eloquently argued that the grouping and metrical rules proposed by the GTTM are present in most African music. However, what is important to note is that the GTTM's rules are open to empirical investigation (granted, however, that there are difficulties involved in quantifying tension and relaxation in prolongation reduction – see Patel, 2008:257-258 for a discussion).

The GTTM has been immensely influential in music cognition research. However, the theory is not without its problems. Some of these problems have been noted by Lerdahl and Jackendoff themselves, and efforts have been made to improve the model. For example, the theory of tonal pitch space (Lerdahl, 2001; TPS) has been proposed to further refine the notions of tension and relaxation in the GTTM's prolongational reduction component (that is, the perception of the ebb and flow of tension in a musical piece). The TPS theory also allows for a quantification of musical tension and relaxation. While claims that tension is perceived hierarchically have been supported by research (such as Lerdahl & Krumhansl, 2007), it is difficult to apply such analyses to music where harmonic tension is not a driving force, but the perception of tension is still generated by other means. An example in this regard is Ravel's famous *Bolero*, which generates tension by means of orchestration techniques and *crescendo*, instead of through hierarchical tonal arrangement.

[100] This despite the fact that Lerdahl discusses the GTTM as a contribution to music theory (see Lerdahl, 2009).

The GTTM does not readily address issues of musical meaning. Of course, the work is a theory about perception of structure, as opposed to the listener's musical experience as a package. This has been emphasised by Jackendoff's comment that he has no objection to the "enhancement of affect associated with activity" being constitutive of musical meaning (Jackendoff, 2009:197). However, if the perception of musical meaning is the result of the causal properties of propositional and non-propositional mental states, as has been argued in this study, it would appear that structure itself has an important role to play in questions of meaning. Theories of musical syntax must seek some sort of connection between syntax and meaning in a more comprehensive manner than in the GTTM. Of course, tension is easily connected to musical affect. Therefore, it would appear that the GTTM formalisation, in combination with the TPS theory, would not require any major theoretical contortions to incorporate an account of meaning (provided that meaning is seen in terms of varying levels of tension, as perceived by the listener). Indeed, such a view may easily concur with the picture that has been painted in this study: syntax resulting in affect.

In this regard, it is worth noting a particularly difficult methodological puzzle concerning the study of the respective syntaxes of music and language. The source of this puzzle is (once again) semantics. Many of the words making up phrases and sentences have arbitrary meanings of their own. It is the meaning of words that helps us to analyse the grammatical structure of a sentence. We have no such helping hand when considering the case of music. The formatives (that is, minimal syntactically functioning units) of language are arbitrary in nature: there is no direct association between the word and its referent beyond the conventions of a linguistic community. The result is that words have a cognitive component that is divorceable from the manner of delivery. As will be argued below, the equivalent of linguistic formatives in music differ in their fundamental nature: they have no dimension of arbitrary meaning. Rather, they depend on their sonic features (including manner of performance) to elicit states in the mind of the listener. The mere fact that words have arbitrary meanings results in grammar being able to influence what is understood by the listener, because by virtue of arbitrariness, syntax can allow for a single word to have multiple meanings. In music, discrete units, such as short phrases or notes, do not have arbitrary meanings. Rather, they act causally. This results in (a) classes of units having one sort of causal affect (for example, loud noises only causing alarm-type states); and (b) a syntax that does not need to unambiguously handle cases where a formative has more than

one meaning. This matter will be further discussed below, as the nature of the formatives of language and music are a major disanalogy that has important consequences for how we think of musical syntax.

6.2 Chomsky's fundamental assumptions and their applicability to the GTTM and cognitive theories

Considering that Lerdahl and Jackendoff have provided one of the most important theories of musical syntax in the form of the GTTM, it would be wise to consider some of the influences brought to the theory by the study of generative grammar in linguistics. In this regard, it is to Chomsky that the authors owe their biggest debt. A cursory reading of the opening chapters of Chomsky's *Aspects of the Theory of Syntax* (1965) – a work with which Lerdahl and Jackendoff were intimately acquainted – provides many basic standpoints on the issue of linguistic syntax that could easily be applied to a study of music.

First, both the GTTM and Chomsky's work use the term 'generative' to describe their respective syntactic theories. Chomsky defines it thus: "by a generative grammar I mean simply a system of rules that in some explicit and well-defined way assigns structural descriptions to sentences" (1965:8). The application of these rules 'generates' a structural description of an utterance. "When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence" (Chomsky, 1965:9). This is not far removed from what we see in Lerdahl and Jackendoff's theory. With the GTTM, the competent listener, upon hearing a musical piece, is able to assign a structural description to the piece – in other words, the listener is able to abstract structural information pertinent to musical understanding. However, Lerdahl notes that whereas Chomsky's model of language starts from an ideal structure (a mental structure) that generates a surface structure (the sentence) through the application of rules, the GTTM starts with a surface structure (the musical piece) and works *backwards* to produce the structure that is heard by the listener (Lerdahl, 2009:188). It is worth pointing out that the reason this particular arrangement was noted was because Lerdahl and Jackendoff could not see how to justify a basic mental blueprint underlying all musical pieces from which all well-formed musical pieces could be produced by the application of generative rules. In Chomsky's theory of 1965, a number of basic strings are assumed to

underlie all well-formed sentences. These basic strings have a structural description that Chomsky calls a "base Phrase-marker" (1965:17ff). Lerdahl and Jackendoff clearly do not make such an assumption with regard to music. In fact, the idea is rejected altogether, along with Schenker's belief in an *Urlinie* underlying all tonal music. Rather, the listener abstracts a structure *from the piece itself*, according to rules. In this way, the GTTM is generative: given a musical piece, the listener generates a structural description according to rules. But whereas Chomsky's transformational model steps off from an ideal structure, to which are applied generative rules in order to produce a surface, the GTTM starts with the surface and arrives at a structural description via the application of generative rules. In Chomsky's model, it is the surface that is generated; in the GTTM, it is the underlying structural description.

The reader should not immediately jump to the conclusion that the GTTM has nothing to tell us about the *production* of musical statements, and that Chomsky's transformational grammar has nothing to tell us about the *apprehension* of sentences. The reason for Lerdahl (2009) explicitly mentioning this distinction has more to do with the sorts of 'mental objects' that are being assumed in either theory. With the theories of Chomsky, many commentators have noted that the cognitive system underlying our linguistic capabilities is constitutive of a mental organ dedicated to language (e.g., Pinker, 1994). Throughout *Aspects of the Theory of Syntax* (and many other works) Chomsky refers to 'knowledge of language' in the mind of the listener. However, it is unclear (and, according to Lerdahl and Jackendoff, unjustifiable) that there is similar 'musical knowledge' in the mind of the listener. That is, of course, excluding the knowledge gained by the competent listener as a result of being immersed in the musical idiom in question.[101] Rather, it is not assumed is that there is an underlying structural description for musical pieces in the listener's mind that is applicable to all idiomatic music, such as knowledge of the *Urlinie* that would be assumed if Schenker's theory were formalised. Such mental objects are the product of generative processes in the GTTM, rather than the start. Structural descriptions are the end product of the application of generative musical rules.

This brings us neatly round to the question of 'rules' in not only the GTTM, but also with regard to our own usage of the term thus far. What are the 'rules' of which grammarians and music theorists speak? When tackling this question, it is important to remember that a theory

[101] Indeed, this is the reason for the proviso mentioned by Lerdahl and Jackendoff that the theory concerns the musical intuitions of persons *experienced* in the tonal genre.

of syntax (such as that forwarded by Chomsky) is a theoretical formalism. Such formalisms have, as their objects, things like 'mental organs' and the 'language faculty' (Hauser, Chomsky and Fitch, 2002). The principles by which things like the language faculty produce everyday language are usually referred to as 'rules', and are the stock in trade of theoretical linguistics. It is primarily for this reason, as well as the use of the term in the GTTM, that this style has been adopted throughout this study. But these rules are part of a formalism, and as such, are merely ways in which we speak of phenomena. Hence, we would not necessarily expect to find a 'rule' hardwired into the brain, or constituting some neural network. However, we may well find a neural network that performs a function that effectively implements a rule we have described in the formalism.[102] Furthermore, it may be the case that a neural network is not dedicated to the implementation of that rule, but simply (by chance) acts in a manner that produces the effect we described with a rule.[103]

Thus, the referent of the term 'rule' is merely a formalised description of a mental phenomenon. The choice of the term 'rule' follows standardised usage in linguistics. In addition, I have been using the term 'well-formed' to designate musical utterances that are in accordance with syntactical rules; this, too, is borrowed from linguistics. Note that these terms carry no connotations beyond their technical usage. The syntactical rules governing the intelligibility of musical (and linguistic) utterances should not be seen as prescriptive rules that somehow impinge on the aesthetic domain of musical experience. A 'well-formed' utterance that is in accordance with syntactical rules is by no means a determinant of aesthetic value, any more than the literary value of a poem or novel is determined by whether the sentences it contains are grammatical or not. Rather, this is a matter of musical cognition and intelligibility.

Lerdahl and Jackendoff formulated two kinds of mental 'rules' to describe the assessment of a piece of music in the tonal idiom. These rules are 'Well-Formedness Rules' and 'Preference Rules'. It is worth exploring the distinction between these two sorts of rules, as they were formulated to address a problem particular to the formalisation of music, and largely absent from language. The reader will recall that it was earlier stated (with regard to syntax and its

[102] An instructive example of how such abstract operations may be realised in the brain can be found with Pinker's discussion of logical operators (1997:98-111).

[103] The reader should note that this discussion mirrors the debate as to whether mental modules have physical reality. Interestingly, two of the most ardent mentalists of recent times, Chomsky and Fodor, argue that mental modules are no more than theoretical abstractions, largely divorced from their physical manifestation. See Fodor (1998) for more on this debate.

implications for the formation of cognitive and affective states) that sounds that could be recognised as musical could be thought of as being placed on a continuum of well-formedness. Some utterances could be understood as music despite featuring several syntactical violations. However, at some point, syntactical violations would outweigh parsable sequences. In such cases, the utterance would not be deemed music at all. Furthermore, there are a large number of utterances that differ with regard to surface structure, but are equally syntactical. The problem is precisely how to formalise such a wide array of possible conditions, *without the help of semantics to guide us in identifying ungrammatical utterances.*

Lerdahl and Jackendoff attempted to solve the problem of distinguishing the array of possible grammatical musical statements by postulating not only rules for well-formedness, but also rules encompassing and differentiating possible realisations of grammatical statements.

> We have found that a generative music theory, unlike a generative linguistic theory, must not only assign structural descriptions to a piece, but must also differentiate them along a scale of coherence, weighting them as more or less "preferred" interpretations (that is, claiming that the experienced listener is more likely to attribute some structures to the music than others)... [they are] *well-formedness rules*, which specify the possible structural descriptions, and *preference rules*, which designate out of the possible structural descriptions those that correspond to experienced listeners' hearings of any particular piece.
>
> (Lerdahl & Jackendoff, 1983:9)

In introducing preference rules, the GTTM is able to describe the myriad of possible fully syntactical utterances that are in agreement with well-formedness rules. It must be noted that no equivalent of preference rules exists in linguistics.

Part of the challenge that linguists face when constructing grammars is accounting for the creative aspect of language, especially in the face of the limited means by which language is implemented. This is the famous 'infinite use of finite means' characteristic of language: with a limited vocabulary and a finite set of rules, language is capable of expressing an unlimited number of thoughts, and every user of language can produce and comprehend sentences

which have never been uttered before.[104] Can a similar thing be said of music? Does music boast this 'creative' dimension? The analogy is, admittedly, a bit strained in this regard. As has been argued at length, the sorts of thoughts expressed in language (namely, propositions) are not communicated via musical means. The question of whether an infinite number of analogous musical thoughts can be communicated seems to be trivial. In any case, the view of communication that was elucidated earlier in this study seems to forgo such discussion of communication somehow transferring information from one mind to another. Rather, it was suggested that it would be profitable to think of musical communication as an attempt to alter the mental states of the listener. If we frame the question about the creativity of music in this manner, we would be able to say that musical statements are able to alter mental states in an infinite number of ways.

Prominent amongst Chomsky's basic assumptions is a distinction between competence and performance. In terms of language, Chomsky considers competence to be the speaker's knowledge of the language (regardless of whether that knowledge is open to conscious introspection or not), while performance is the use of language in everyday situations (1965:4). This distinction is often coupled with the notion that linguistic theory deals with an "ideal speaker-listener", that is able to wield language in a manner not limited by non-grammatical constraints, such as memory limitations and mistakes in application of linguistic knowledge (p. 3). The purpose of this ideal speaker-listener assumption has to do with the investigation of the mental capacity for language in isolation from factors that otherwise disrupt performance. While the performance of speakers is a part of linguistics, it is seen as a topic distinct from the construction of grammars.[105] It could be noted, however, that competence is not reflected perfectly in performance in the real world, but only in the ideal conditions assumed for the study of grammars. That is, performance only reflects competence perfectly when the speaker-listener is the idealised one assumed by grammarians for the purpose of studying the computational procedures underlying language (Chomsky, 1965:4).

Do we need to make such a distinction between competence and performance for our theory of music in the mind? It is unlikely. Like Lerdahl and Jackendoff, I find it difficult to justify the idea that there is musical knowledge in the mind that is applied to musical pieces, a sort

[104] Chomsky attributes this idea to Wilhelm von Humboldt in the Preface to *Aspects of the Theory of Syntax* (1965:v).
[105] Despite this, Chomsky suggests that worthwhile theories of performance all take ideal grammars as a starting point. See Chomsky (1965:10ff).

of structural yardstick by which musical pieces are judged (that is, serving as the 'competence' part of the dichotomy). This is the view suggested by Raffman (1988; 1993) when she suggests that there is a musical schemata in the mind to which incoming musical pieces are compared. Of course, I am not denying that musical knowledge can be built up over time. But I do not see sufficient reason to posit a mental schema or idealised structure for musical pieces in the mind, somehow analogous to Chomsky's "base Phrase-markers" – a limited subset of ideal musical strings in the mind of the listener. It is this justification that Lerdahl and Jackendoff also found to be problematic. It underlies the notion of the GTTM working backwards, as discussed above. That is, whereas language involves a structural description that produces a surface (the sentence) through the application of generative rules, musical cognition starts with the surface and then uses rules to generate a structural description. I believe that in the case of language, if structural descriptions are such an important part of language usage, they may (in some sense) be innate or somehow facilitated by mechanisms in the mind of infants to aid with the acquisition of language. It would appear that a similar argument for innateness with regard to musical abilities is at odds with current evidence (Patel, 2008a:377-385). That is, the case for innate linguistic structural descriptions appears to be much more plausible than the idea of innate musical structural descriptions. Rather, it may be that the rule system for music apprehension is innate. In other words, cognitive processes handy for the generation of musical structural descriptions may well be innate, but the structural descriptions themselves by no means are. This is an important point to keep in mind when approaching the syntax of music from the vantage point of linguistic syntax: whereas a case can be made for innate structural descriptions in language, the same cannot be said for music.

A major component of Chomsky's theories about language is the notion of deep and surface structure. The deep structure of a sentence is related to the surface structure by means of transformational rules. Chomsky notes that, with regard to a sentence, it is "a *deep structure* that determines its semantic interpretation and a *surface structure* that determines its phonetic interpretation" (1965:16; Chomsky's emphasis). For example, it is a common deep structure that results in the following four sentences meaning roughly the same thing, despite having different surface structures (example after Aitchison, 1989:98):

1. Charles captured a rabbit.
2. A rabbit was captured by Charles.

3. It was a rabbit which Charles captured.
4. What Charles captured was a rabbit.

These four sentences share a deep structure, which determines the semantic interpretation of each sentence (and ensures that a competent English speaker's interpretation of each is basically the same). They differ in terms of the means by which they are realised – that is, they differ in terms of surface structure, their phonetic interpretation. In Chomsky (1965, 1972), it is held that the deep structure and the surface structure are linked by 'grammatical transformations': "rules expressing the relation of deep and surface structure" (Chomsky, 1972:166; as quoted in Aitchison, 1989:98).

Is it profitable to think of music in terms of deep and surface structure? First, it would be important to again remind ourselves that there is no semantic component to music. Therefore, any notion of a musical deep structure as related to semantics must be modified. We might imagine that a sense of deep structure could be understood if one speaks of either elaborations of a basic musical phrase (similar to an *Urlinie*), or possibly the manner in which thematic material is changed or developed in the course of a musical utterance. Let us examine this question with regard to the GTTM. Lerdahl and Jackendoff rule out the notion of deep structure from their own formalisation (Lerdahl & Jackendoff, 1983:278ff). However, they do not write deep structure off altogether. Rather, they suggest that what may be of interest are "simple, normative, archetypical forms" underlying more complex surface structures (p. 288). They speculate that there may be a sense of 'transformational rules' in virtue of which the possible surface structures that may be generated from an archetypical form. However, they are hesitant to refer to such forms as deep structures (*ibid.*). Instead, they suggest that such archetypical forms (if they exist) are ultimately the products of the preference rules expounded in the GTTM.

The problem with thinking about musical syntax in terms of deep structure is two-fold. First, music has no semantic component. Deep structure was posited as an explanation for the phenomenon of sentences with differing surface structures but similar meanings. Music, it would appear, has no such bedrock of semantic meaning. Without semantics, it is difficult to see what precisely the purpose of a deep structure for musical utterances is. Second, the notion of deep structure assumes mental objects in the mind that serve as the elementary

statements. I concur with Lerdahl (2009) that such mental objects are difficult to justify in the case of music.

A salient point of difference between music and language that must be reflected in any theory of musical syntax is the fact that in music, multiple utterances can be combined in a manner that retains (to a large degree) intelligibility. These are cases of voice-leading and polyphonic texture, characteristic of many sorts of music in the world. This is also a facet of musical syntax that is not handled in the GTTM or any other cognitively-orientated theory of which I am aware. This has already been alluded to as the perceived multi-dimensionality of musical experience. Music is not only perceived in terms of time (that is, horizontally), it is also perceived in terms of space – that is, horizontally, with more than one event being apprehended at any given moment. This scenario is not optimal for effective language use, as it would hinder the ultimate purpose of propositional communication. It is vital that in language performance, semantic meaning is communicated as clearly as possible. With musical utterances, however, multiple events can (and do) contribute to the formation of meaning-relevant mental states. Future theories of musical syntax must confront this feature of musical experience, and recognise it as a major point of difference between music and language.

Part of the reason why the generative theories of Chomsky and other linguists of the 1950s and 1960s were seen as revolutionary has to do with the fact that universal claims about the nature of linguistic behaviour were made. An important objective of modern theoretical linguistics has been describing those features of languages and linguistic behaviour that are common to all humans. These features underwrite and partially determine the nature of the various cultural manifestations of language. Instead of a focus on the phenomenon of language as a particular, localized phenomenon, Chomsky and his followers sought to describe languages in terms of their commonalities (or, more accurately, in terms of universal cognitive features reflected in the output of a hypothesised language faculty). The question of universality with regard to musical behaviour has already been discussed at length. The GTTM project, too, made certain claims with regard to universality, despite the fact that the study was focused on tonal music. However, much empirical work still needs to be conducted before the validity of Lerdahl and Jackendoff's claims can be properly ascertained.

6.3 Groups as formatives in musical syntax: a disanalogy between music and language

A fundamental difference between music and language is the nature of the constituent units making up utterances. This is particularly noticeable when music is compared to the elements in language which we will refer to as 'formatives'. This word has been chosen after Chomsky's definition, where formatives are seen as "minimal syntactically functioning units" (Chomsky & Halle, 1965:98; see also Chomsky, 1965:3). (For a description of formatives in a more modern and technical linguistic sense, see Mohanan, 1995:40-43. I will be using Chomsky's definition.) Superficially, what Chomsky is referring to are those bits of language that are combined and manipulated by syntactical means: words, affixes, suffixes, and so on (that is to say, not only words, but also affixes and suffixes such as 'de-', 're-', 'ex-', '-ed', and '-s', which are subject to description in syntactic terms). Below the level of the formative, sounds have no meaning whatsoever and can make no syntactical contribution to linguistic utterances. It is formatives that are governed by grammatical rules.

Can we say that there are units similar to Chomsky's formatives in musical syntax? In other words, are there minimal syntactically functioning units in musical utterances, units upon which syntactical principles act? It will be argued here that there are. Not to argue for minimal syntactically functioning units in music would beg the question of what it is that our postulated musical syntax governs, because to assume a musical syntax is to assume that there are units that the syntax regulates. It will also be argued that the nature of these minimal syntactically functioning units (musical formatives) is radically different from those of language. This difference principally stems from the fact that formatives in language are essentially arbitrary symbols that act in a cognitive manner, whereas those of music act causally. The formatives of language have conventional meanings, and linguistic syntax governs these units in virtue of their semantic qualities. Of course, a suffix such as '-ed' does not have a referent, but its appendage to a word such as 'smoke' can be syntactically described and has an impact on semantics. Thus, in language, we can say that formatives are 'semantically rich'. According to the model of musical meaning discussed in this study, we could say that the formatives of language have the power to elicit cognitive mental states by virtue of their function as both syntactic and semantic units.

Musical formatives, on the other hand, do not feature a semantic dimension. Rather, they have an impact on mental states by directly causal means. To appreciate this point fully, we

will have to consider in more detail what exactly constitutes a musical formative. The definition cited above calls for a 'minimal' syntactically functioning unit. In the case of language, this can be taken to refer to the smallest unit that can be said to have any meaning at all, or whose combination with another unit results in a change in meaning for that unit. Minimal units are not, however, the only parts of linguistic utterances upon which syntax has an impact. Units above the level of the formative are also affected by syntax, for example noun and verb phrases. In this manner, we can say that formatives are the bottom in a hierarchy of syntactically relevant groups, up to the level of a sentence (i.e., formatives and words, to phrases, to sentences). It seems to make sense to say that music operates in a similar manner. Single sound events are at the bottom of the hierarchy; what follows are motives, then phrases, and so on. That is, musical formatives (single musical discrete elements) are the things that are minimal syntactical functioning units.

What are examples of single musical discrete elements or musical formatives? The most obvious candidate would be a single note – a sound with a clearly definable fundamental pitch and a duration, which is discrete.[106] However, further reflection will reveal that a single note cannot serve this function adequately in light of music's lack of semantics. With language, any sound can serve as a formative: all a linguistic community needs to do is agree on what that formative refers to, or what syntactical role it is chosen to fulfil. What separates formatives from other discrete ambient sounds is the fact that an arbitrary meaning has been assigned to the formative. Thus, a drum strike can count as a formative in English, with a syntactic function, if a community agrees on a meaning for the sound (and, consequently, a syntactic function – for example, the drum strike could be a name, and fulfil the function that regular nouns do; alternatively, it could be used as a verb, with similar syntactic results). Arguably, certain suffixes and affixes that are combined with other formatives are semantically based in some sense, as they always affect the semantics of other formatives in the same manner. For example, despite not having a fixed referent, the affix 'non-' implies negation, as in 'non-governmental'; '-ex' is usually associated with 'former' or 'no longer', as in 'ex-communist'.[107]

[106] The property of discreteness means that there is no more sense in speaking of 'half a note' then speaking of 'half a suffix'.

[107] This is not a hard and fast rule. It is arguable that the suffix '-al' has no particular sense, but does act syntactically in rendering words in their adjective form, e.g. 'fundamental', 'syntactical', and 'formal'. A similar example is the past tense function of '-ed', as in 'worked', 'played', and 'faded'.

However, the notes of musical utterances have no such semantic foundation, nor can they be used predictably to change syntactical function. An E flat in isolation is not identifiable in terms of its potential function in a musical phrase (if it is considered in isolation). Furthermore, a single E flat is indistinguishable as a musical note unless it is embedded in a musical phrase with other notes. Without an arbitrarily assigned meaning, an E flat may as well be simply another ambient sound that happens to have a discernable pitch. Only when the E flat is in combination with another pitch, and a relationship between the pitches is established, is a musical formative in evidence. It is the relationship between two musical discrete elements that constitutes a musical formative. Therefore, the minimal syntactically functioning units of music must consist of at least two musical discrete elements, so that a syntactical relationship is at least implied. And, whereas a linguistic formative can elicit cognitive states in virtue of its semantic qualities, a musical formative can only act causally, and hence result in non-propositional affective states.[108]

The issue of minimal syntactically functioning units can be described in terms of the theory of cognitive states that has been the focus of this study. Linguistic formatives function as syntactical units in virtue of their impact on semantics. The result of semantics on mental states is of a cognitive nature: words, for example, are represented in the mind as cognitive states. Cognitive states are propositional in nature, and are achieved by non-directly causal means. Musical sounds (indeed, the remainder of sounds) achieve mental states largely due to causal relationships, where the sound as it is immediately represented in the mind can be seen as being in a causal relationship with regard to the resultant non-propositional mental state. This is the difference between musical formatives and formatives in language. In language, minimal syntactically functioning units have a cognitive dimension, and in music minimal syntactically functioning units have a causal dimension. Thus, the formatives of language can result in cognitive states, whereas musical formatives cannot, and instead operate principally by virtue of a causal relationship between the sound and a resultant non-propositional mental state.

[108] This discussion of musical formatives can be fruitfully compared to the notion of 'musemes' forwarded by Tagg (1979), and discussed by Middelton (1990). Designed to be directly analogous to the morpheme in linguistics, the museme is the most elementary figure that can be said to have musical meaning. However, as our discussion thus far has dealt with the notion of musical meaning, to adopt the concept of a museme would imply a hint of circularity.

6.4 Grouping

It is argued here that the musical equivalent of linguistic formatives is largely a perceptual matter (as opposed to a propositionally-processed cognitive matter), and that musical perception will have a large part to play in any formulation of musical syntax. Whereas a large portion of linguistic syntax depends on semantics in some manner, musical syntax has far greater concern with aural perceptive principles in general. Therefore, many aspects of musical syntax are not music-specific in any meaningful way. Such a view is congruent with the theory sketched in this study, namely that a large part of the apprehension of musical meaning has to do with the causal powers of ambient sounds in eliciting non-propositional mental states. With natural language, both the meaning of words and the structure of the language's grammar provide a means of interpreting the otherwise meaningless strings of noises emitted by another person. Linguistic syntax, in particular, provides a means of demarcating and grouping semantic structures and relations, which are then apprehended as compound but discrete semantic entities (e.g. phrases). Without structural rules for understanding grammar, the identification of sentences (the statements of a language) is not possible. The application of grammatical rules to a sequence of sounds, in combination with the assigning of a meaning to the individual elements within the sequence, allows us to distinguish communicative statements from other ambient sounds. In order for meaning to be generated by structure and semantics, a communicative medium must have elements upon which semantic and syntactical rules can operate: in other words, parsing of auditory input in the form of discrete units is fundamental to mental processing with respect to communicative acts.

It is broadly conceivable that musical groups play an analogous *functional* role to words in language (that is, as discrete entities), but it is not entirely clear that they satisfy the 'minimal' condition for formatives (that is, 'groups' are conceivably reducible to smaller groups, and eventually to single musical events – it is a minimum of two musical discrete elements that play the role of formatives in music). Of course, the syntax of language and the syntax of music allow for additional generation of meaning over and above the individual meaning of discrete units or formatives. Any attempt to draw an analogy between groups and formatives in music to words or phrases and formatives in language will bring to light an interesting feature characteristic of musical formatives. Whereas the formatives of language are (in some cases) combined with other formatives to construct words (such as the combination of *work*

and *-ed*, or *non-* and *entity*), their hierarchality tends to end there. Clarity and brevity are, after all, important and essential requirements of natural language.[109] With musical formatives, however, the hierarchical possibilities are much further exploited. A single musical formative, such as D-A motive, is usually compounded with many others to create phrase-length musical utterances. Whereas a typical English sentence may have around a dozen to two dozen formatives, a typical musical statement of comparable length may have dozens more. A single chord opening a romantic piano sonata may contain upwards of five notes. In Western polyphonic music before the 18th century, the issue becomes even more marked. In such music, each voice has a melodic identity in addition to its contribution to the harmonic progression present in other similarly handled voices. Therefore, the listener is confronted with a far greater number of minimal syntactically functioning units than with a linguistic utterance, which is obviously possible since musical formatives do not require any semantic effort on behalf of the listener.

What then, constitutes a group in musical utterances, and what is their position in musical syntax? Lerdahl and Jackendoff take grouping perception to be hierarchical all the way up, from a single event, into groups (perceived as such by virtue of Gestalt-like perceptual principles), to phrases, sections, through to entire musical pieces. The rule stating that "a piece constitutes a group" (GWFR 2; Lerdahl & Jackendoff, 1983:37-38) shows the extent to which they project hierarchical grouping. However, there is little made in the GTTM of the fact that Gestalt grouping rules operate not on the level of the musical piece, but rather within smaller, localised time spans. That something such as the grouping of similar sections (e.g. a repeat of an exposition in sonata form) is not to be explained in terms of Gestalt principles describing the grouping of similar notes occurring after one another, is not a seriously discussed issue in the GTTM. It is unlikely that any perceptually-derived rules operating on localised events (such as the first three notes in the famous opening of Beethoven's Fifth Symphony) can be described as operating on events occurring at a much higher level (such as the repeat of the exposition in the same movement). A formalisation of musical syntax will have to mark the distinction between events occurring on the level of phrases and those occurring on considerably larger levels (such as sections). Formalising this is not an easy

[109] Theoretically, the combination of formatives to create words need not be constrained by the demands of efficiency. Some monstrous words can be created, for example 'antidisestablishmentarianism'. However, in practice, everyone but bureaucratic tyrants and academics avoid such words. Many styles of music, however, thrive on such rich combinatorial potential, made possible by the lack of a need to unambiguously communicate propositions. Also note that linguistic syntax permits infinitely long sentences because of recursion.

problem to solve. Theoretically, then, grouping is hierarchical, as Lerdahl and Jackendoff suggest; in terms of the types of rules that act at the level of the phrase and at higher levels in the hierarchy, it would appear that this is not necessarily the case.

Grouping, nevertheless, plays an important part in our apprehension of musical statements. The primacy of grouping in musical syntax is helpfully explained by Lerdahl and Jackendoff:

> The process of grouping is common to many areas of human cognition. If confronted with a series of elements or a sequence of events, a person spontaneously segments or "chunks" the elements or events into groups of some kind. The ease or difficulty with which he performs this operation depends on how well the intrinsic organization of the input matches his internal, unconscious principles for constructing groupings. For music the input is the raw sequences of pitches, attack points, durations, dynamics, and timbres in a heard piece. When a listener has construed a grouping structure for a piece, he has gone a long way toward "making sense" of the piece: he knows what the units are, and which units belong together and which do not. This knowledge in turn becomes an important input for his constructing other, more complicated kinds of musical structure. Thus grouping can be viewed as the most basic component of musical understanding.
>
> (Lerdahl & Jackendoff, 1983:13)

It can be taken as given that musical phrases are perceived as groups, themselves being constructed hierarchically. Thus, there is a hierarchical organisation *within* the phrase. We are able to distinguish groups of notes within phrases, and those groups of notes can be further separated into distinct, single minimal syntactically functioning units (musical formatives), and at the lowest level, single discrete musical events. It is worth noting that in many sorts of human music, it is possible to syntactically alter intra-phrase groups and still arrive at intelligible phrases. For example, it is possible to imagine many phrases of Viennese classical music (for example) with different opening motives, provided that syntactical determinants such as tonality are preserved. We can imagine a piece of music having intra-phrase groups substituted with altered ones; as long as the modifications are syntactical, the phrase is perceived as musical.

There is a fair amount of consensus that music is temporally perceived in terms of grouping. Furthermore, it would appear that much of this grouping adheres to the same sorts of principles described by the Gestalt psychologists of the early 20th century, such as Wertheimer (1923). The reliance on Gestalt perceptive principles is present in much work on music perception, including work by Lerdahl and Jackendoff (1981; 1983; 2006), Shepard (1964; 2001), Deutsch (1980; 1982; 1992), Tenney and Polansky (1980), Levitin (2006), and Temperley (2000), to name but a few. It is important to realise that the majority of Gestalt perception principles formulated in the early years of Gestalt psychology were geared toward the explanation of visual grouping. Shepard (2001:32) provides some relevant Gestalt grouping principles, applicable to visual processing, that easily find correlates in aural perception:

a) Proximity: Objects that are close together are perceived as belonging to the same group.
b) Similarity: If objects are equally spaced, those that are similar are grouped together (provided they are adjacent).
c) Symmetry: Symmetrical patterns are grouped together, due to the fact that it is statistically unusual to find a degree of symmetry amongst objects arranged randomly.
d) Good continuation: collinear objects tend to be grouped together.
e) Common fate: objects that move together are usually perceived as a single group.

These principles may seem self-evident or even trivial. However, the fact that these principles dominate our day-to-day interaction with external phenomena is important. There are numerous examples of objects which have no real intrinsic or causal connection being perceived as forming a single group. These cases of false correlation show how dominant our perception of grouping is.

To the above must be added the intuitive concept underlying most Gestalt theory, namely the idea that groups are seen as more than the sum of their individual parts. A visually perceived group is more than just an assemblage of parts; a forest is more than a thousand individual trees, and a rabbit is more than simply "undetached rabbit parts" (to paraphrase Quine). This is likely to go far toward explaining why we tend to hear chords not as collections of different

notes, but as units that function in a particular manner and have characteristics that can be attributed to the chord as a whole. It is also this intuition which is probably at the base of Lerdahl and Jackendoff's 'Grouping Well-Formedness Rule 2' – namely, that a piece constitutes a group, and that "a piece is heard as a whole rather than merely as a sequence of events" (1983:37).

Before continuing, we may be prompted to pose an interesting question regarding Gestalt-like perception: to what extent would evidence of the same grouping principles applying to non-musical aural stimuli be an indication that Gestalt perception in music is the result of a general brain mechanism? Does it just happen to be that evolutionary pressures hit upon the same grouping principles for the perception of both our visual and our aural environments, and that both use the same system? Is it most plausible to say that music-related grouping perception is parasitic on our general aural grouping capabilities? (Of course, Gestalt-grouping is a very successful cognitive shortcut – for purposes of survival, it could hardly be a hindrance.) The answers to these questions undoubtedly lie in neuroscience, and it is an answer which I am not suitably qualified to give. However, aural and visual grouping principles, if as important to survival as they appear to be, were probably in place long before the rise of the hominid line, and will be present in many non-human animal perceptive systems. If fact, we would expect them to have been in place for a very long time indeed: probably as long as hearing and vision have been a dominant factor in the survival of complex organisms. Therefore, I think that it is fair to assume that Gestalt-grouping in music is a function of general perceptual systems that apply equally to all aural perception. However, the ability to detach or abstract such groups from their general auditory context, and then consider them in relation to other similarly detached groups according to rules, is a human-specific feature of auditory cognition.

How do these grouping principles manifest themselves in musical perception in particular? Tenney and Polansky (1980) see the temporal Gestalt units of musical perception as being built up from small elements that are not temporally divisible, such as single notes.[110] These temporally indivisible units form, in turn, larger and larger groups in hierarchical fashion (1980:207). Further support for the idea that the basic grouping principles present in music

[110] A single note is, for Tenney and Polansky's (and our) purposes, temporally indivisible in that it makes little sense to speak of 'half a note', or to say that we perceive the first note as a group made up of two instances of 'half-the-first-note'. In other words, temporally indivisible single musical elements are what we have thus far been referring to as discrete.

are similar to Gestalt principles comes from psychologist Diana Deutsch. In her 1980 article, Deutsch suggests that a fundamental principle of musical processing is the grouping of music, usually along the lines of Gestalt principles (pp. 168-170). Despite the age of this article, it shows some good examples of musical grouping principles in action. Furthermore, Deutsch suggests that these grouping principles sometimes 'override' other parts of our processing. For example, in fast pseudopolyphonic scale runs (leaping monophonic lines with alternating high and low notes, with the high notes descending and the low notes ascending), perception of irregularities in time between the two lines is less successful than in a single voice, single register pattern (p. 171). This strongly suggests a preference to group similar registers together, as well as providing an interesting counter-example to Lerdahl and Jackendoff's contention that only temporally adjacent events can form groups. Other perceptual grouping phenomena are described by Deutsch, such as those that arise from pseudopolyphonic passages that are heard through headphones, during which the stereo effect is changed by panning from one ear to the other (illustrated in Deutsch, 1980:172) – yet another example of perception being governed by grouping principles rather than remaining faithful to the stimulus. This is what we would expect if grouping principles were fundamental to musical and aural processing.

Another Gestalt grouping principle is that of common fate. Shepard (2001) expands on the principle of common fate, noting that "[d]emonstrations of auditory common fate typically involve common onset time, common amplitude modulation, and common frequency modulation" (2001:33). Shepard states that this principle underlies our ability to distinguish between different instruments in (for example) an orchestral setting. According to Shepard, it is the grouping of harmonics according to the time of onset, common amplitude, and presumably related frequencies that allow us to tell one instrumental event from another. This is also possible if the fundamental pitches of two instruments are the same: difference in timbre from two different sources (such as a trumpet and a cello) allow us to discriminate according to the principle of common fate. This is mirrored by Levitin's description of the grouping of complex overtones into the sound of a single trumpet playing a single note (2006:77).[111] Thus, it would appear that the discrimination of discrete pitches from one another is aided by principles that could be partially explained by auditory scene analysis (as described in Pinker, 1997:529-539). To what extent these are instances of the same level of

[111] In general, Levitin's remarks around (2006:76-80) are the same as those expressed by Shepard (2001). There is little discussion of the grouping of multiple notes into in motives and phrases.

operation of Gestalt principles as discussed by Lerdahl and Jackendoff is a moot point, as these are cases of the perception of single musical events and not syntactically operating portions of utterances. In terms of the structure of music (i.e. how it is organized), Lerdahl and Jackendoff have more to say.

Lerdahl and Jackendoff view grouping and metrical structure as separate sets of rules (1981:480), although one set may reinforce perception of the other. One example is the musical phenomenon of pulses, or beats. At this point, it is important to distinguish between a recurring pulse or beat, and the perception of a hierarchical meter. The ability to entrain to an external beat has not been observed in animals that have a close phylogenetic relation to humans, and along with the interactive behavioural aspect of entrainment, is a possible candidate for a music-specific evolutionary adaptation (an opinion shared by Patel, 2008a:402ff). Metrical hierarchy, however, is possibly a factor in the grouping of sounds during language acquisition in 7 month old infants (Hannon & Johnson, 2005; cited in Trehub & Hannon, 2006), although these infants cannot yet entrain to an external beat (Patel, 2008a:405). Thus, it may be that the metrical perception of regular pulse is just an instance of a general perceptive principle concerned with the hierarchy of events in time (Patel, 2008a:406).[112] However, it is the metrical hierarchy of a regular pulse that Lerdahl and Jackendoff are interested in. Regarding this, there are a few subtle points made: (a) beats are points in time, and do not have a duration – they are hypothetical; (b) the periods of time that pass between the occurrence of beats are called time-spans; (c) beats are always perceived as regular (Lerdahl & Jackendoff, 1981:489).

Lerdahl and Jackendoff arrive at the following conclusion regarding the nature of groups and metrical structure: "grouping structure consists of units [i.e. physical bits of music] organised hierarchically; metrical structure consists of beats organized hierarchically" (1981:494). Many of these units are instances of aural Gestalt perception. Lerdahl and Jackendoff consider grouping principles to be the most basic and fundamental elements of "musical understanding" (p. 481). A possible conclusion to be abstracted from Lerdahl and Jackendoff's text is that grouping involves partitioning the actual musical sounds into groups based on cues from the music itself, whereas metrical hierarchy is something that we perceive

[112] It is on this basis that the work of Trehub and Hannon (2006) is called into question by Patel (2008a:406). He similarly cites the conclusions of Hannon and Johnson (2005) and Phillips-Silver and Trainor (2005) as problematic with regard to the nature of beat perception in infants below the age of 4.

due to the pulse and events on the surface of the music. Grouping is not strictly hierarchal in nature, because groups can overlap (elide) on multiple levels (see Lerdahl & Jackendoff, 1981:481-482; this was a principle determining factor in the proposal of what they call 'transformational rules' – not the same sort of transformational rules as can be found in generative grammar, however).[113] Grouping is recursive in the sense that groups can be part of other larger groups: "Because of this uniformity from level to level, we can assert that grouping structure is *recursive*, that is, capable of indefinite elaboration by the same rules" (Lerdahl & Jackendoff, 1981:483). Grouping structure is perceived as more fundamental than the stresses of the meter, i.e. grouping principles tend to override metrical perception. According to Lerdahl and Jackendoff, this is evidenced in music such as Gregorian chant and the *alap* introductions of the North Indian raga (p. 486), where grouping is favoured over adherence to strong and weak beats in meter.

It would be worthwhile to note that Lerdahl and Jackendoff distinguish three types of 'accent' present in the musical utterance that aid with our abstraction of the metrical hierarchy. An event that occurs on the foreground or surface of the musical line, that gives emphasis or stress to the particular moment of that event, is called a *phenomenal accent* (Lerdahl & Jackendoff, 1981:485). For example, articulation that functions as an emphasis of a note can be classed as a phenomenal accent. It is worth noting that phenomenal accents would fall into the class of phenomena we called 'surface variables' in the previous chapter. *Structural accents* are accents affected by virtue of harmonic or syntactic stress (*ibid.*), regardless of other events that bring further attention to that moment (for example, a phenomenal accent of some kind). Structural accents are caused by the rules of the GTTM. *Metrical accent* is different in that it is supplied by the listener. This is the 'perception' – or more accurately, imposing of – beats and a beat hierarchy in a piece of music. This is supplied with the help of cues from the music itself (that is, via the other rules of GTTM). Once metrical accent has been inferred, it is not easily abandoned by the listener, unless there is strong evidence to the contrary (see for example Temperley, 2000). Syncopation occurs when the conflicting evidence on the musical surface is not strong enough to affect a change in metrical accent in the listener's perception (Lerdahl & Jackendoff, 1981:485-486; see also the discussion of syncopation in Temperley, 2000:82-86).

[113] Note that the relation between non-adjacent but similar groups (e.g. group b in [a a b a a b]) is governed not by grouping rules, but by rules of associational structure (1981:484). In other words, grouping the two instance of b together is not done in virtue of Gestalt rules, but rather due to association. This type of structure is not hierarchical.

The rules for grouping and rhythm proposed by Lerdahl and Jackendoff (1981; 1983:345-347) have been taken up by ethnomusicologist David Temperley. In an article on African rhythmic grouping and meter (2000), Temperley finds that broadly construed, the metrical and grouping preference rules from GTTM apply to the African music that he investigated. Temperley concludes that the grouping principles for Western music (that is, the principles expounded by Lerdahl and Jackendoff) and those for African music are very similar. This is what we would expect to find if fundamental Gestalt-type perceptive principles were instrumental to grouping processing, as they are common to the human species. Indeed, Temperley notes Lerdahl and Jackendoff's belief that the grouping principles "reflect 'gestalt' [sic] principles of similarity and proximity which are known to apply to perception in general" (Temperley, 2000:90; citing Lerdahl and Jackendoff, 1983:40-43).

As might be expected, Temperley found that the perceptive phenomenon of meter extraction was present in the African music he surveyed. Often, it is found that this underlying pulse is not actively expressed in the music itself. Instead it is expected that the listener supplies the pulse. The generally accepted tempo range of 80-170 b.p.m. for the tactus is regarded as common ground between Western and African music.[114] If the tactus is inferred from music which does not readily supply obvious cues from which to abstract the tactus, we may expect to find that less obvious cognitive rules serve as a guide to the listener and performer. An example is the type of continuous syncopation that we find in African music, as well as the use of stressed syllables in sung lines over weak beats in much African music. Furthermore, just as we find with Western music, the African music surveyed by Temperley suggests that there are only duple and triple relationships between one rhythmic level and the next level up in the hierarchy (2000:20), a feature claimed to be characteristic of Western tonal music by Lerdahl and Jackendoff.[115]

[114] The tempo range of 80-170 b.p.m. for dotted crotchets is, according to Temperley, fairly close to the range that Lerdahl and Jackendoff propose as the ideal tactus in Western music (1983:73; as cited in Temperley, 2000:69).

[115] Interestingly, Lerdahl and Jackendoff consider this feature as *not* universal (1983:69). The sort of music that may have prompted them to claim this is some varieties of Balkan folk music (such as the Slovenian music cited in Omerzel-Terlep, 2000:913). For example, meters such as 5/8 are perceived in terms of *non-regular* beat lengths, such as 2+3 and 3+2. However, I have been unable to ascertain whether there is any 'naturally' occurring music that features groups of 2+2+1 for 5/8 time. What I am implying here is that I think it may be so that beats are never perceived in groups of 1 (i.e., two strong beats cannot be adjacent at any level).

Patel, in mentioning that both language and music boast rhythmic organisation of some kind, suggests that perceptual grouping is a major intersection between the two phenomena (2006:99). He claims that music and speech both feature rhythmic perceptual grouping: "the mental clustering of events into units (e.g., phrases) at different hierarchical levels" (2006:99). He cites evidence from neuroimaging suggesting that grouping perception in language and music require the same hardware in the brain (*ibid*.; Knösche et al, 2005; Patel, Peretz, Tramo & Labrecque, 1998). Patel writes that "grouping in music may well be an offshoot of prosodic grouping abilities" (p. 99), a hypothesis worth investigating.[116]

By way of comparison, Trehub and Hannon (2006) define grouping, rhythm, and meter as falling under the general topic of temporal processing (pp. 81-82): "*Grouping* refers to the perception of boundaries between groups and subgroups of elements in an unfolding musical sequence. *Rhythm* refers to the pattern of temporal intervals in a sequence… *Meter* refers to the abstract, hierarchical structure of music, which is experienced as a strong and weak alternation of 'beats'" (Trehub & Hannon, 2006:82; note that Trehub and Hannon subsume both meter and pulse under the heading 'meter'). They appear to suggest that musical grouping principles observed in pre-linguistic infants[117] are the same as those used as cues in the appropriation of environmental stimuli for language acquisition. This would suggest that group cue awareness is not music-specific, but rather a part of language acquisition. Furthermore, it is suggested that the use grouping principles is not species-specific either (p. 83). The link between grouping and language acquisition is further supported by Patel (2006; 2008a). All of this in turn supports our hypothesis that grouping principles are general cognitive mechanisms for the perception of utterances of both music and language.

As a side note: rhythmic processing of some form is also found by Trehub and Hannon to be used for language cues, as well as not being human-specific. Meter is linked by them to the ability to synchronise oneself with music. They state that no activity in a non-human animal has been observed with regard to music, although they suggest that meter may be the product of general mechanisms for cyclical patterns in organisms the world over. Basically, Trehub and Hannon are wary of saying that music or any of its constituents is modular (domain-specific), especially because of the links with language learning and the rhythmic grouping

[116] Patel suggests that there may well be a hierarchy of meter in speech, but not a regular pulse (2006:100).
[117] Trehub and Hannon are not very clear as to precisely what these grouping principles are. However, it seems to be suggested that they are falls in contour toward the end of groups – pauses perceived as being in the middle of such falls are generally perceived with alarm.

abilities of non-human animals. This runs contrary to the view expressed by Peretz and Coltheart (2003), who saw fit to describe some aspects of music cognition in terms of domain specificity (i.e. dedicated processing evolved for musical activities).

6.5 The question of narrative or coherence rules above the level of the phrase

The majority of our focus thus far has been on musical utterances that have a temporal duration roughly equal to that of sentences, given physical restraints such as working memory. It was mentioned that the principles in operation at the level of the phrase cannot be assumed to be the same principles that govern larger structures. That is, it cannot be assumed that phrase-relevant rules are simply reapplied at higher hierarchical levels to handle larger musical structures. This is analogous to the idea that grammatical rules work on the level of sentences, but do not apply to entire paragraphs (groups of sentences). The principal reason why this is not possible for the case of music is because the phrase-relevant rules are derived from localised perceptive principles that are not in operation over longer periods of time. What we should now consider is the idea that there is a distinct set of principles operating at these higher hierarchical levels, ignoring for now the precise circumstances under which these higher-level rules come into effect. This is similar to the idea of some sort of set of principles governing the understanding of the logical course of argument in a group of sentences.

In the case of music, let us refer to this possible set of rules as *narrative* or *coherence* rules.[118] These are the sorts of principles that govern how phrases themselves follow one another, how several phrases combine to form a recognisable section of a piece, all the way up to the level of the piece as a whole. What should be immediately apparent is that these principles – if they can be said to exist – are of a far weaker variety than the perceptive principles operating at the level of the phrase. They should be thought of as analogous to the sorts of guiding principles for constructing arguments in language, as if there was some sort of implication of logic guiding the manner in which certain arguments are made. This point is illustrated by

[118] 'Coherence' is chosen after 'discourse coherence' (Patel, 2008a:335-342; Kehler, 2002). I have added 'narrative', as 'coherence' is already a music-theoretical concept regarding thematic material in musical pieces. However, 'narrative' has also been a concept used to describe the course of musical pieces within the New Musicology paradigm (see, for instance, Almén, 2008). The usage here has only a superficial similarity to that found in 'narratological' studies in New Musicology.

Patel (2008a:335-342), in a discussion on 'discourse coherence', drawing upon the work of Kehler (2002).[119] Consider the following examples (from Patel, a8:335):

1. The father saw his son pick up his toy chainsaw. Seashells are often shiny on the inside. John likes peas.
2. The father saw his son pick up his toy chainsaw. The boy pretended to cut down a tree, but didn't touch the delicate flowers growing in the garden. Mom was pleased.

The sentences in the two examples quoted above are grammatically well-formed, providing no violations of standard English grammar. However, a listener is likely to judge the sentences of (1) as having no relation to one another – as being incoherent as an argument or piece of discourse. The sentences comprising (2) are, however, coherent. Hearing (2) as coherent requires several assumptions. For example, the father's son and 'the boy' are understood as referring to the same person; it must be assumed that what Mom was pleased about was the fact that the boy meant no ill-harm to her flowers; and that the make-believe cutting was done by the toy chainsaw. In the examples stated above, it is notable that discourse coherence depends greatly on semantics. It is the meanings of the sentences that allow the second and last sentences of (2) to follow each other coherently. This is a major obstacle to the application of Kehler's principles to music, as music has no semantic component to aid coherence. It is clear then that if coherence principles are in operation in music, they would not be operational by virtue of semantics. What then results in us hearing two phrases as part of the same piece of music, or a phrase entailing another phrase? Or, in virtue of what do we judge successive phrases and sections to be parts of the same musical pieces?

In the case of Western tonal music, tonality often serves as a determinant of coherence. For instance, if confronted with a 16-bar section in a piece of Viennese Classical music with a tonality of C major, comprised of an 8-bar forephrase and an 8-bar afterphrase, the piece would be considered less coherent if the afterphrase were transposed to a remote key, such as F sharp major. Similarly, if the afterphrase was completely atonal, it is unlikely that an

[119] The nature of discourse coherence is discussed in more detail by Patel (2008:335-342).

unprepared listener would consider it to be part of the same piece as the forephrase. However, Western music's fixation with tonality is mostly culturally determined. Many sorts of music do not have a similar emphasis on tonal function and the idea of transposable tonal centres. If we imagine a music that is mono-tonal (that is, remaining in one key), how is coherence further achieved? It may be the case that two successive sections have a common tempo, a common meter, or a similarity in thematic material. Successive sections may have similar pitch contours, or similar rhythmic treatment of material. Timbre and register may be other factors, as can less precise similarities, such as similarity in articulation. These factors lend a degree of coherence to successive sections. The judgement of coherence will be based on a (possibly weighted) judgement of the degree of similarity of successive sections in comparison with differences. Judgements of coherence in musical structures larger than a phrase are likely to place much reliance on statistical processes. For example, successive sections may have a common meter, but different timbres, rhythmic treatment, contour, and tempo, resulting in a lack of coherence between these sections. Coherence rules are unlikely to be hard and fast. Instead they will have to entertain a measure of flexibility, possibly being similar in nature to the preference rules described in the GTTM.

If coherence rules exist and can be described as such, it is likely that a large proportion of them will be stylistically (and hence culturally) determined. For example, sections with different tempos that are unrelated thematically are perceived as separate sections, and were it not for their conventional placement alongside one another, would not be thought of as belonging to the same piece of music. In a similar manner, we are able to distinguish between the movements of a symphony; and despite the fact that the movements of early Classical symphonies are related tonally, we judge them to be separate pieces of music on the basis of tempo and thematic content. Some factors contributing to coherence, like tempo and meter, register, melodic contour, timbre, rhythmic identity, and possibly tonality in some weak sense, may not be culture-specific. I feel that there is an argument to be made for the idea that some basic features of coherence are the product of perceptive discrimination. While the task may appear daunting, it is easily imaginable that matters of coherence are empirically testable in discrimination tests. Thus, researchers may be able to determine what factors lead to a decrease in coherence between successive sections by asking test subjects to discriminate between sections on basis of perceived coherence (e.g. as to whether adjacent sections are considered part of the same or different musical pieces).

Whether coherence rules that are not culturally-derived are the result of general principles of cognition or of the physical constraints of the matter of the mind is an open question. However, it should be noted that in general, as we move from phrases to larger structures, the amount of mental, biologically determined rules progressively diminishes, and the importance of conventional, culturally determined factors increases.[120]

6.6 Overview: Basis of a theory of musical syntax

Let us consider where this leaves us. A large portion of this study has been devoted to pointing out that because of music's lack of semantics, any theory of syntax will ultimately have to incorporate an account of how it is that music is often perceived as meaningful. Whereas semantics provides a bedrock of meaning that aids the theorist in the construction of linguistic grammars, music has no such semantic helping hand. Furthermore, in the mental picture of meaning that has been sketched in this study, the manner in which music is put together is an important contributor to its meaning, as syntactical features contribute directly to the formation of (especially affective) mental states. The syntax of language does contribute to meaning in some sense, in that the ordering of constituent elements of sentences has an impact on meaning. But the syntax of language does not, to any sizeable extent, contribute to the formation of affective states. The syntax of language functions as a facilitator of cognitive states. Therefore, future theories of musical syntax cannot afford to discriminate between matters of meaning and matters of structure.

An adequate theory of syntax will also have to account for what has been termed the vertical dimension of musical utterances: the fact that in music, there is often more than one discrete event occurring at any given moment (for example, in polyphonic or chordal passages, where many notes or melodies occur at one time). This is not a major concern for linguistic syntax. The goal of syntax in language is to aid the relay semantic meaning as unambiguously as possible. Considering that spoken language is most efficient when there is only a single horizontal line of discrete elements being uttered at any given time, it is not surprising that linguistic syntax has no means for coping with multiple events occurring at once. However,

[120] The term 'biologically determined' includes what might be considered 'psychologically determined' rules, thus accounting for the idea that some rules, and the perception of some features of musical utterances (such as perceived tension and resolution, or tonal closure) are better characterised as psychological (as opposed to biological in the hard-wired, genetically-determined sense).

the vertical arrangement of simultaneous events happens in many sorts of music. A theory of syntax has to be able to handle this vertical dimension of music: the fact that multiple lines can be combined to form intelligible utterances, with more than one discrete event occurring at any given moment.

How would an ideal theory of syntax look? The theory will have to clearly identify what sonic elements are being ordered by syntactical rules. Therefore, a minimal syntactically functioning unit, or musical formative, will be the unit upon which syntactical principles operate. It was argued here that because of the lack of arbitrary meaning being attributed to single musical events, a minimum of two musical discrete elements should count as a musical formative in virtue of the fact that some syntactical function is implied. Single musical events considered in isolation (e.g. a single sound with a discernible fundamental frequency and a duration) are indistinguishable from ambient sounds unless they are heard in relation to other events, implying or fulfilling some musical-syntactic function. Formatives are combined into groups by virtue of Gestalt-like perceptive principles, and these groups are combined to create phrases in a hierarchical manner.

A distinction between phrases that are comparable in length to spoken sentences, and larger scale structures, needs to be made by any syntactical theory. A theory will also have to account for the perception of musical events on levels higher than the phrase, as well as explaining how coherence between adjacent phrases is achieved. In music-theoretic terms, it seems to make sense to say that phrases are hierarchically combined to create larger-scale structures such as sections through to entire pieces. Lerdahl and Jackendoff advocate a similar scenario for grouping: from groups of notes, through to phrases, sections, and eventually entire pieces. However, the grouping principles that regulate grouping up till the level of the phrase are unlikely to be the same as those governing the perception of sections or pieces as groups, due to limitations on our immediate cognitive grouping abilities. Hence, we would expect that grouping perception is at its strongest within the temporal duration of the phrase, and becomes weaker as the temporal duration increases.

The various principles making up the syntactical organisation of phrases is likely to have a degree of biological dependency (such as a dependency on Gestalt grouping principles). However, as we proceed from groups and phrases to supra-phrase structures (such as forephrase and afterphrase sections, up to larger sections and eventually pieces) the

dependency on biological principles is likely to progressively diminish. It still appears that at supra-phrase levels, a certain amount of coherence is at play in musical pieces. A theory of syntax will need to describe how this sense of supra-phrase coherence comes about, as well as distinguishing coherence rules from the rules that operate at the level of the phrase. As the temporal duration of the structures in question increases, the role of biological syntactical determinants decreases, and culture-specific factors become more pronounced. Therefore, due to the universality of cognitive processing, one would expect to find most syntactical universals on the level of the phrase (such as Gestalt-type grouping principles), with more cultural specificity on higher hierarchical levels.

A further question that will ultimately have to be confronted by theorists attempting to formalise the syntax of musical cognition is whether the physical manifestation of syntax in the mind is by nature cognitivist or physicalist. In other words, are the mental rules present in musical cognition the result of general mental principles, or are do they reflect physical limitations on mental computation? This is a complex question which waits upon empirical evidence from cognitive science and neuroscience before it can be answered in full. However, it is worth noting that Patel's 'shared syntactic integration resource hypothesis' suggests that there are aspects of cognition shared by both music and language processing (2008a:183-298). This would imply that there are aspects of musical syntax that are not music-specific. However, whether these syntactical resources are the ultimate product of the manner in which cognition happens to work (cognitivist), or more fundamental physical limitations on cognition arising due to the matter and construction of the brain (physicalist), remains a challenging question.

CHAPTER 7
CONCLUSION

7.1 Introduction

> Now suppose that [Lerdahl and Jackendoff] are right; does it follow that music is a language? That is not a meaningful question because the notion of what is a language is not a meaningful notion. Is it a human language? No, of course not; it is not a human language. Is it *like* human language? Well, sure, in some ways but then the question is: how 'like' do you mean?
>
> To say that something is a language is not a meaningful comment; it is just to say – 'it is enough like human language so that I'll call it "language".' It is like asking: does somebody live near Boston? There is no definite answer to that.
>
> (Chomsky, 2000:44-45)

Language has acted as the analogical foil against which we have considered the nature of musical meaning. Music and language are obviously not different versions of the same phenomenon; one is not just a wholesale specialisation of features found in the other. However, they are not completely isolated behaviours either, whether in biological or conceptual terms. As was suggested in the second chapter of this study, it is helpful to think of music and language as communicative media that operate in a similar way to writing, intentional gestures, and other symbolic and iconic activities; in that the primary objective of such behaviour is to alter the mental states of conspecifics. If communicative media are thought of in this manner, one can then imagine that music and language are cases of communicative media where one or another particular function has been specialised. Even then, however, differences exist. Semantics is the prime example. In cases where music can be said to be referring by the use of conventional, arbitrary symbols, it has been suggested that music is simply acting as language. (Although, obviously, the use of words in combination with a purpose-built computational system is generally more efficacious than

leitmotivs.) In such cases, regular semantics is in operation, and meaning can be studied in the same way that it is studied in cases of semantics in linguistic behaviour.

Even though much research still needs to be conducted, there seems to be little doubt that in cognitive terms, music and language are related. What remains to be seen is in what further respects and to what extent music and language share neurological realisation, and what implications such knowledge will have for our conception of the independence and domain-specificity of either behaviour. In conceptual terms, there seems to be much to learn in pursuing an analogy between the two behaviours. Scruton holds that analogies are "based in resemblance and they are illuminating only if the resemblance is deep, so that the knowledge of the one thing casts light upon the other" (1997:171). Of course, pursuing an analogy that is not entirely apt may prove an equally profitable exercise in conceptual terms. Deep differences in the music-language analogy may help us to understand why they are separate phenomena; when the analogy breaks down, we have a critical angle at which to investigate differences between the two phenomena. Semantics was a case in point. Absolute music per definition does not refer to things beyond music. When music does come to act semantically, it is functioning like language. Semantics is a feature of language that is simply not present in music, until music is treated as having the linguistic property of semantics. (That is, there are no separate semantic systems for language and music respectively in the brain. Rather, there is a single semantic system, which is exploited when any communicative medium refers to objects and concepts beyond itself. Of course, the principal *raison d'être* for the computational system Language [the faculty of language in the narrow sense, FLN], when utilised for communication, is the reliable implementation of semantics.)

Apart from the matter of semantics, it would appear that, when regarded as communicative media, music and language share many mechanisms for eliciting mental states in the minds of listeners. Both result in cognitive and affective states via the manner of their performance, and hence trigger processes in the mind of both a computational and directly causal variety. Although the study was geared toward explaining music, the application of these ideas to studies of linguistic semantics and the mind, as well as other sorts of communicative behaviour, is possible. This broader understanding of 'meaning' can encompass both semantic and pragmatic considerations in language- and music-related meaning. As I am aware of no such similar approach to the topic of musical meaning, it would be wise to

consider some of the ways in which the approach of this study differs from other theories of musical meaning.

7.2 Key differences in approach

7.2.1 *The term 'semantics'*

As this study draws much of its inspiration from analytic philosophy, clarity with regard to the term 'semantics' has been an important point. We utilised a definition of Tarski (1944) to define the term accurately. According to this definition (SDS; Ch. 3), semantics is the study of the relations between propositions and the states of affairs in the world which they can be said to refer to. This implies that reference is a necessary prerequisite for semantics; hence, any behaviour that is used to communicate without propositions cannot be said to have a semantic dimension. This narrow definition is congruent with the modern analytic philosophical usage of the term 'semantics'. A hallmark of propositions is truth and falsity: a proposition can be said to be true or false when compared to the state of affairs in the world that it refers to. Truth and falsity do not apply to non-programmatic music. A minuet cannot be 'true' or 'false', simply because it does not refer to states of affairs beyond the music. Music cannot therefore be thought of has having a semantic dimension.

The model advocated in this study differs from other approaches to meaning in music in that it adheres to the philosophical definition of 'semantics'. This is principally a means of avoiding confusion. The established currency of the term can lead to confusion when it is loosely employed with regard to both communicative meaning and meaning in its non-communicative sense (as something's value or significance for someone; see Ch. 1). For example, the manner in which semantics is discussed in Kramer (2002; see especially pp. 151-159) leaves the reader to assume that music does indeed refer arbitrarily to things beyond music in a manner similar to words in language. But Kramer blurs the distinction between meaning as the property of a communicative act, and meaning as the significance of music to someone. A similar scenario exists in the semiological work of Nattiez (1990; 2004). By adhering to established use of the term in analytic philosophy, a field where communicative meaning is a major topic of investigation, this study avoids confusion by clearly marking the distinction between semantics and propositions on the one hand, and the various mental states

that result in the apprehension of what we call 'meaning' in communication in general, all the while not confusing the communicative and non-communicative senses of the word 'meaning'.

7.2.2 *Ontological questions*

Questions of ontological status have been major concerns in music philosophy (e.g. Ingarden, 1986; Kivy, 2002; Davies, 2003; amongst others), with many opposing views being supported. The philosophical webs that are spun in such arguments can be rather thick. For instance, there are those that hold that the musical work is an abstract object, somehow existing in some Platonic sense beyond the physical confines of time and space. In an article entitled "Musical works as eternal types", Julian Dodd claims that there is no relation of identity between an entity called the musical 'work' and performances of the work (Dodd, 2000). In other words, the 'work' is an object existing independently of performances of it. Works, says Dodd, are "abstract objects which have sound-sequence-occurrences as instances" (p. 242). This places him in rough alignment with the theoretical view of Levinson (1990), who, according to Dodd, argues that musical works are types whose tokens are instances (performances) of them. The argument becomes irretrievably complex when Dodd considers an argument discussed by Levinson. Levinson notes that if works are abstract 'sound structures' that are independent of time and space, then it would be impossible for them to be created by composers. What follows is a complex discussion of how to adequately qualify musical works in Platonic terms, by suggesting an alternative understanding of the notion of 'creation' with regard to musical composition.

The only possible question of importance with regard to the ontology of musical works to issues of meaning seems to me to be that of which identity-granting parts of a musical work are responsible for the elicitation of mental states, and which are dependent on aspects of performance. In other words, are there aspects of instances of similar musical pieces which boast features that result in the perception of the music as meaningful, that can be attributed to all instances of that musical piece? And are there meaning-relevant features unique to performance? However, on the view of musical meaning sketched here, the question of the ontology of musical works seems to me to be largely irrelevant to musical meaning. As the phenomenon of musical meaning is a mental affair, we would expect that if the musical

'work' – whatever it is – has some contribution to the perception of meaning, that is to be explained in mental terms. In other words, even if a work has a reality independent of the listener, be it nominal, Platonic, or whatever, it makes no difference to meaning, as the listener only perceives what she *hears* as a single instance of a work (or more accurately, the mental representation of what is heard). Of course, past experience of similar performances of a musical work have an impact on what we ultimately come to understand as the work's significance for us, and we can become deaf to certain features which upon earlier listening may have resulted in affective states.[121] Of course, a listener can be conditioned to react differently to a musical work upon repeated listening. But the features of some abstract entity are not an issue here: rather, it is the similarity of repeated stimuli that leads to such conditioning. Thus, this model simply sidesteps the issue of ontology because it is perceptually irrelevant.

7.2.3 *Avoidance of arguments claiming that meaning is the perception of some specific feature of musical experience*

The discussion of meaning as it is constituted in the mind by means of mental states allows us to avoid the tricky situation where we would have to contend that the perception of some particular aspect of the music *is* meaning (e.g., meaning is understanding the formal qualities of the music). Rather, it is suggested here that musical meaning is a mental state that is *caused*, either computationally or causally, by relevant physical properties or features of musical utterances. It is the perception of those aspects of musical utterances that results in the mental states that we in turn perceive as musical meaning. This means that we do not need to say that 'the perception of musical meaning is the perception of *x*'. Rather, we are able to say 'the perception of musical meaning *results from* the perception of *x*, by definable processes', where *x* constitutes meaning relevant properties of musical utterances.

The problem with views that musical meaning is the perception of some or other feature of music is that such views are limited in their explanatory power in only being able to account for meaning in music that boasts that feature. For example, it is commonplace for authors to

[121] Of course, a fundamental assumption in this study was naturalism akin to the causal closure of physics. Therefore, any view of musical ontology would consequently have to be of a naturalist variety. The idea of a Platonic ideal of the musical work existing independent of space and time must therefore be rejected on grounds of the causal closure of physics.

claim that the meaning of music has to do with emotion (Brown, 2000; Mithen, 2005), or, more broadly, 'feelings' (Raffman, 1993; Jackendoff, 2009). Apart from the general vagueness of many of these explanations (for instance, Brown's assertion that the meaning of music is simply emotion), the implication of such views is that music that is not seen as emotional, such as Venda children's music or modern electronic dance music, is considered to be meaningless in the communicative sense. Furthermore, such views presume that there is something intrinsic to the structure of emotionally-relevant music that is necessarily absent from music without that structure, with the latter music being meaningless.

7.2.4 *Universality and ethnocentricity*

The model forwarded here is intended to be construed broadly enough for it to be applicable to a wide variety of musical styles, genres and cultures. The argument from mental states is widely applicable, and is not restricted to any particular genre of music. Many other sorts of models of musical meaning are genre-dependent. Semiotic approaches (such as Nattiez, 1990, 2004; Tarasti, 1995) depend on the nature of the musical 'signs' themselves, and are hence genre-specific. Cooke's *The Language of Music* (1959) depends on melodic and harmonic figures from tonal Western art music of the common practice period. Other approaches, such as that of Kramer (2002), reduce matters of musical meaning to social context, and are hence also culture-specific.

Having been inspired by recent research suggesting that human musicality is biologically determined, this study has sought to steer clear of the twin troubles of ethnocentrism and presentism (Cross, 1999; 2001a). Ethnocentrism is here taken to be an erroneous assumption that conclusions about the object of study in one's own culture apply equally to the same object in other cultures, without due consideration of their universal validity. Presentism is the chronological equivalent: the assumption that the object, as observed in the present, is identical to the object as it was in the past. It is believed that these two assumptions have, as far as possible, been avoided in this study. In other words, the theory forwarded here is stated in terms general enough to have applicability to music as it is practised now in diverse cultures, as well as in the past.

An example of a view of music that is considered by Cross (1999) to be ethnocentric is the well-known passage on music's status as an evolutionary adaptation in Pinker (1997:529-539). Pinker suggests that music is a prime example of a universal human activity that is nevertheless not an evolutionary adaptation. Instead, it is the sort of activity that merely 'tickles' the reward systems of the brain, in the same way that we have a tendency to indulge in sweet and fatty foods because of in-built reward systems that provided an incentive to find these rare resources in the past. Pinker went as far as to claim that music was "auditory cheesecake" and would not even be missed by humans if it were not there, in the way in which cheesecake has not been seen as instrumental to human survival. After an initial backlash from the music-psychological community (see Trehub, 2003; Cross, 1999, 2001a; Huron, 2003; Levitin, 2006; for example), it has generally been held that this view may well be at least partly true. (It is a conclusion arguably forced by Patel's recent consideration of music as an adaptation [Patel, 2008a].) In the dozen or so years that have passed since Pinker's pronouncement, the idea that musical behaviour is, in part, parasitic on other aspects of cognition has generally come to be accepted because of the view that 'musicality' consists of a package of abilities, some of which are not music-specific.[122] However, it is the means and not the validity of Pinker's final conclusion that are of interest to us here. According to Cross, "all the evidence that Pinker presents and the assertions that he makes as to the nature of music in human experience are again most directly applicable to what music has become over the last hundred years within technologised and capitalistic Western society: an aural commodity to be consumed, dispensable on command" (Cross, 1999:7). In other words, the idea that music is something that is made by a specialist class of people in society with special skills, and passively listened to by the rest for pleasure only without any physical involvement, is not a condition of music the world over. Rather, it is simply the manner in which it occurs in the industrialised and commercially-driven Western world wherefrom Pinker hails; a condition of listening to music that has been shaped by the powers of radio and the recording industry over the past century. It is therefore not indicative of all the world's music. A similar failure with regard to ethnocentrism is that of Miller (2000), who backs up his argument of an origin for musical behaviour in sexual selection (specifically mate attraction by males) by citing infamous cases of modern rock musicians' promiscuity. Miller has mistaken a single form of music, where the specialist performer is idolised by

[122] For example, one can see an early attempt at a modular interpretation of musical cognition in Peretz and Coltheart (2003). An adapted version and discussion of the conclusions reached in that paper can be found in Mithen (2005).

listeners, and whose music is often distributed *en masse*, and used this as evidence that musicality in humans developed from courtship rituals. However, few societies elevate their musicians to the status of superstars, and few societies draw a distinction between musicians and listeners as separate social classes (Cross, 2001a). Therefore, the case of the sexual prowess of a rock star in Western popular culture cannot be attributed to musicians of other cultures.

We have already seen some of this ethnocentrism occurring with regard to the dominance of the relation of performances to the 'work' concept in the philosophy of music. The interest in the concept of the work, with single authorship, and its reception by audiences, has resulted in an interest in its relation to individual performances. However, as mentioned above, the idea of the work-concept has limited applicability to the world's music. Any Platonic idea of an ideal work surely cannot apply to spontaneous improvisation or the trance music of the San people. The problem of ethnocentrism is most pronounced when one considers the formulation of musical syntax, as it should be universally applicable. Hence, at the more elementary levels of musical syntax described in the previous two chapters, it was suggested that general and universal cognitive principles play an important part; and as the variables multiply in longer musical utterances, the role of contextual and cultural parameters in basic musical intelligibility becomes more pronounced.

7.3 Is 'meaning' still missing from the computational machine?

The argument thus far can be summarised as follows: when somebody perceives a piece of music as meaningful, what occurs in the mind of the listener is a change in mental states. This change is essentially caused by the musical stimulus, as it is immediately represented in the mind of the listener. There are two types of mental state, cognitive and affective. When somebody perceives a piece of music as meaningful, all the mental states (which are arrangements of matter in the brain) can be described in terms of these two sorts of states. A change in mental states reflects a change in the arrangement of matter in the brain. These states are, in some way, constitutive of the meaning that is perceived by the listener. When someone truthfully says, "I perceive this musical piece as meaningful", they are actually saying that the piece of music in question has caused certain sorts of mental states to exist in

their mind. Thus, 'meaning' is a mental phenomenon, and never independent of minds; however, it is determined by relevant causal characteristics of the stimulus.

This appears to leave us with a scenario where the description of mental states exhausts the explanation of musical meaning. In other words, there seems to be nothing more to be said about musical meaning other than mental states and the processes of causality responsible for their existence. 'Meaning' would simply be a way we speak about these brain states, just like 'heat' is how we speak of the energy of particles, or 'sound' is used to describe the compression of particles in air. In a sense, we would have 'reduced' the phenomenon of musical meaning to simpler things (mental states). This raises the possible criticism that, by reducing a complex, subjective phenomenon to more basic components that are not individually instances of meaning, we have an explanation of meaning with no meaning in sight. It would appear that we have described a mechanistic system in the mind, and meaning – the phenomenon that we had originally wished to explain – is nowhere to be found. By explaining meaning as mechanistically as our description of mental states, it would appear that there is no longer any meaning in our answer! We could refer to this intuitively postulated missing meaning as the 'meaning missing from the computational machine', the computational machine being the description of mental states as constitutive of meaning.[123]

There appear to be two ways out of this intuitive impasse that leave the argument intact: (a) by suggesting that the intuition that meaning is missing from the explanation is fallacious by virtue of circularity. We want to explain meaning, and if there is still meaning in our answer, then the argument must be circular. This is akin to asking what ingredients are to be found in a particular cake, and arriving at an answer of eggs, milk, flour, and cake. In this case, missing meaning is nothing more than an illusion, as mental states exhaust the explanation of meaning; (b) the intuitive meaning missing from our account can be explained as something else that is not mechanistically (that is, causally) connected to the processes described in this study. For example, a possible suggestion is that the missing 'meaning' in the computational machine of the mind is actually a dimension of the actual qualitative experience of those mental states. Thus, when we look at the constitution of meaning in the mind from a third-person perspective, our intuitive feeling that meaning has not been explained can be

[123] This term is chosen after the term 'the ghost in the machine', which was used by Ryle (1949) to describe the Cartesian notion that within the material body, a non-material spirit is to be found, despite the problems of how material and non-material entities could possibly interact. Ryle argued that Descartes had arrived at this conclusion by fallacious thinking (in particular, category mistakes).

accounted for in that it is only with the first-person experience of those states that meaning achieves its subjectivity. The meaning that is felt to be missing from our otherwise mechanistic explanation would then lie in the subjective, first-person experience of those mental states. This would also explain why our third-person mechanistic explanation seems to lack this intuitive missing meaning.

Let us consider case (b) further. Any attempt to account for the intuition of missing meaningfulness by appealing to our subjective, qualitative experience of meaning, is essentially explaining the experience of meaning in terms of *qualia*. Qualia are the subjective experiences in consciousness: the smell of coffee, the redness of colour, the taste of wine. These are very real parts of our everyday experience of the world – it's the way things seem to us, from the subjective, first-person vantage point. Usually, issues surrounding qualia arise in explanations of consciousness. They are of bearing to our discussion in that case (b) is an appeal to the subjective nature of the experience of mental states – what it is like, from a first-person perspective, to entertain meaning-relevant mental states. Ultimately, one must decide whether this experience can be explained physically, or whether it in some sense lies beyond physics. For example, it is easy to explain all the physical facts about tasting wine, but not as easy to explain *what it is like* to taste wine. If a wine scientist knew absolutely everything about the physical causes and effects of tasting wine, but had never drank wine himself, would he learn something new when he tasted it for the first time? If your answer is yes – he does learn something new – then you must conclude qualia to be non-physical, because our scientist knew all the physics already. But if you answer no – he already knew what wine tasted like subjectively before drinking it (by virtue of his omniscience regarding the physical facts of wine-tasting) – you are suggesting that the subjective experience of qualia can be accounted for physically.[124]

The problem raised concerned the intuition that something is missing from the theory of mental states being constitutive of meaning (or meaning being reduced to mental states). Objection (b) implies that a purely mechanistic explanation of the mental processes occurring in the minds of listeners fails to account for the subjective experience of meaning. Arguments about the nature of qualia become relevant in the following scenarios:

[124] Example after Jackson (1982).

1. The ‘meaning missing from the computational machine’ is our qualitative experience of affective and cognitive mental states, but meaning itself is constituted, indeed exhausted, by mental states and the processes that determine them. What is intuitively felt to be missing is the qualitative experience of these states, and not the determinants and mental constitution of meaning itself. Therefore, our explanation of meaning is complete if it accounts for the mechanistic principles underlying the formation of mental states, and any sense of missing meaning is simply a case of mistaking the experience of meaning-relevant mental states for the meaning-constitutive mental states themselves.
2. The ‘meaning missing from the computational machine’ is our qualitative experience of affective and cognitive mental states, and it is this qualitative experience that *is* meaning. Mental states and procedures are just mental states and procedures; meaning itself is the qualitative experience – the qualia – of entertaining those states in your mind.

(1) reduces meaning to mental states, and suggests that the ‘missing meaning’ is mental state qualia. With (1), the intuition that meaning is missing from our explanation is simply a case of mistaking meaning itself for the experience of meaning. (2) suggests it is the quale itself that is the meaning (including missing meaning). In (2), meaning is exhausted by what we experience in first-person perspective, and the mechanistic explanation deals with the machinery behind the experience. In both instances, the missing element is not denied, but is as real (and as difficult to explain) as the qualitative experience of the taste of coffee or the redness of red. Both suggest that the qualitative experience is missing from the theory; but (1) says that ‘missing meaning’ is just qualitative experience and not meaning; and (2) says the qualitative experience of mental states *is* meaning.

This is a subtle difference and I am not sure if the argument could be settled. According to both (1) and (2), the ‘meaning missing from the computational machine’ has something to do with qualia. With (1), the intuitive ‘missing meaning’ is simply the experience of mental states. With (2), meaning in its entirety is the experience of mental states (that is, mental states and processes are nothing more than mental states and processes, and ‘meaning’ is the experience of being subjected to those states and processes). So, with (2), what seems to be the case is that whatever one’s position on the fundamental nature of qualia is, will determine the ultimate nature of this missing meaning. If you take qualia to be non-physical, then

according to (2), the missing meaning will be non-physical and irreducible to a mechanistic explanation. With (1), the missing meaning is not equivalent to qualia, so the nature of qualia does not determine the ultimate nature of meaning.

Alternatively, the objection that there is something more to meaning that simply explaining mental states can be approached from another angle. When confronted with the intuition that the explanation does not account for the phenomenon of musical meaning in its entirety, we can interrogate ourselves as to precisely what we are looking for when we claim that something is missing. Although it is tempting to call this missing element 'meaning', it may not be meaning at all. Our intuitions may be satisfied if we look for some other element of musical experience that is distinct from communicative meaning. For example, it may be that we feel that there was some more abstract goal behind the composer or performer's choice of musical utterance: the composer's ultimate intentions. A composer may use an utterance that strikes us as sad or forlorn. In and of itself, this is all that is communicated. However, if understood with the composer's ultimate intention in mind – say, using the utterance to express grief at the loss of a loved one – the utterance may be afforded a further sense of profundity. It is this sort of deeper intention behind musical statements that no mechanistic theory of mental states can fully account for, simply because it lies beyond the explanatory boundary of communicative meaning heard in isolation from context.

There is, I think, another similar issue to keep in mind, which was initially raised in Ch. 1. The intuitive 'missing meaning' should not be confused with the value we attach to musical meaning; that is, a value judgement *about* meaning, that is not itself constitutive of meaning. That is, it is possible that the missing meaning is just the value we place in communicatively meaningful music. Consider language: statements in conversation and statements in Shakespeare are both meaningful in a communicative sense, and the meaning in Shakespeare can be explained by means of cognitive and affective states like any other linguistic statements. It is the *value* which we place in the mental states that Shakespeare conjures up that makes it so significant. As another example, consider *Finnegan's Wake* by James Joyce. The sentences are non-syntactical and difficult to read. However, it is the value that we place in *Finnegan's Wake* as a major literary artwork that results in its special meaning *for* us, which is something different from the meaning of the statements themselves. Phrased differently, two senses of the word 'meaning' must not be confused: meaning as a property of communication ("the meaning of a sentence or musical statement") and the meaning of things

in our lives ("the meaning of James Joyce", or rephrased, "The significance of James Joyce to me or to society").[125]

7.4 Conclusion

In the introductory chapter, it was mentioned that this study found its principal inspiration in recent work on the biology and evolutionary history of human musicality. It has been proposed that one of the most vexing features of music – that it is often perceived as being meaningful in some ambiguous sense – can be explained as a mental phenomenon arising from our experience of musical stimuli. At this level of conceptual analysis, music and language seem to have much in common. Like spoken language, music can be used to communicate in virtue of the causal properties of utterances. This conceptual agreement is assuring in the light of recent empirical work that suggests that parts of our musical and linguistic processing are interlinked. If music and language to some extent share a biological realisation, we would expect them to have an intertwined evolutionary history. Keeping this in mind, questions of evolutionary history may be informed by a conceptual analysis that sees music as a potential communicative medium with the property of communicative meaning. One would expect that if music and language are to be explained in mutual evolutionary terms, the nature of their functioning as communicative media would need to be explained with reference to both behaviours.

The human species is a musical one. All over the world, large resources of time and energy are spent on the making of music for pleasure, ceremonies and rituals, religious purposes, social reasons, and even economic gain. Because music keeps us very busy, it comes as no surprise that there is intense interest in how it manifests itself in our everyday lives. This, added to the fact that music and musical experience is so highly valued by communities, may partly account for the fact that the problem of musical meaning has become a major topic of interest. However, the problem is not intractable: with adequate conceptual clarity, much of the mystery surrounding musical meaning and its role in musical experience becomes subject to investigation. With advances in our understanding of the biological relation between music and language processing, the problem of musical meaning appears to be a problem of

[125] That is, meaning as a property of communicative acts versus "meaning *of* something *for* somebody".

understanding our communicative apparatus and behaviours in their entirety. Pre-empted by the GTTM, Raffman's attempt to define ineffable musical experience in terms of mental computation provided an important example of how to approach a subjective phenomenon of musical experience that was previously taken as a mysterious fact of musical life. Instead of merely accepting that there was something unexplainable about our attachment to music, Raffman sought to clarify the issues surrounding musical ineffability. I hope that in this study, a similar aim has been achieved: further clarification of a subjective phenomenon, that is now becoming a point of interest in investigations into the fundamental nature of musical behaviour.

REFERENCES

Almén, B. 2008. *A Theory of Musical Narrative*. Bloomington: Indiana University Press.

Aitchison, J. 1989. *The Articulate Mammal: An Introduction to Psycholinguistics*, third edition. London: Unwin Hyman.

Bernstein, L. 1976. *The Unanswered Question: Six Talks at Harvard.* Cambridge, Mass.: Harvard University Press.

Bickerton, D. 1996. *Language and Human Behaviour*. Seattle: Washington University Press.

Blacking, J. 1995. *Music, Culture, and Experience: Selected Papers of John Blacking*, ed. R. Byron. Chicago: University of Chicago Press.

Blacking, J. 1976. *How Musical is Man?* London: Faber and Faber.

Blacking, J. 1967. *Venda Children's Songs: A Study in Ethnomusicological Analysis*. Johannesburg: Witwatersrand University Press.

Blackmore, S. (ed.) 2006. *Conversations on Consciousness: What the best minds think about the brain, free will, and what it means to be human.* Oxford: Oxford University Press.

Botha, R. 2000. Discussing the Evolution of the Assorted Beasts Called 'Language'. *Language and Communication*, 20 (2): 149-160.

Brown, S. 2000. 'The "Musilanguage" Model of Music Evolution.' In N. L. Wallin, B. Merker, and S. Brown (eds.), *The Origins of Music*, pp. 271-300. Cambridge, MA: MIT Press.

Brown, S., B. Merker, and N. L. Wallin. 2000. 'An Introduction to Evolutionary Musicology.' In N. L. Wallin, B. Merker, and S. Brown (eds.), *The Origins of Music*, pp. 3-24. Cambridge, MA: MIT Press.

Burkholder, J. P. 2006. 'A Simple Model for Associative Musical Meaning.' In B. Almén and E. Pearsall (eds.), *Approaches to Meaning in Music*, pp. 76-106. Bloomington: Indiana University Press.

Carston, R. 2008. Linguistic communication and the semantics/pragmatics distinction. *Synthese*, 165: 321-345.

Chalmers, D. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Chomsky, N. 2000. *The Architecture of Language*, ed. N. Mukherji, B. N. Patnaik, and R. K. Agnihotri. New Dehli: Oxford University Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 1988. *Language and Problems of Knowledge: The Managua Lectures.* Cambridge, MA: MIT Press.

Chomsky, N. 1972. *Problems of Knowledge and Freedom*. London: Fontana.

Chomsky, N. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, N., and Halle, M. 1965. Some Controversial Questions in Phonological Theory. *Journal of Linguistics*, 1 (2): 97-138.

Clarke, E. 2005. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. New York: Oxford University Press.

Cook, N. 2001a. Theorizing Musical Meaning. *Music Theory Spectrum*, 23 (2): 170-195.

Cook, N. 2001b. ‘Analysing Performance, Performing Analysis.’ In N. Cook and M. Everist (eds.), *Rethinking Music*, pp. 239-261. New York: Oxford University Press

Cook, N. 1987. The Perception of Large-Scale Tonal Closure. *Music Perception*, 5 (2): 197-205.

Cooke, D. 1959. *The Language of Music*. Oxford: Oxford University Press.

Cross, I. 2005. ‘Music and Meaning, Ambiguity and Evolution.’ Preprint version. In D. Miell, R. MacDonald, and D. Hargreaves (eds.), *Musical Communication*, pp. 27-43. Oxford: Oxford University Press.

Cross, I. 2001a. Music, science, and post-positivist pluralities. *Intégral*, 14/15: 20-26.

Cross, I. 2001b. Music, cognition, culture and evolution. *Annals of the New York Academy of Sciences*, 930: 28-42.

Cross, I. 1999. ‘Is music the most important thing we ever did? Music, development and evolution.’ In S. W. Yi (ed.), *Music, Mind and Science*, pp. 10-39. Seoul: Seoul National University Press.

Cross, I., & Tolbert, E. 2009. ‘Music and meaning.’ In S. Hallam, I. Cross and M. Thaut (eds.), *The Oxford Handbook of Music Psychology*, pp. 24-34. Oxford: Oxford University Press.

Cross, I., & Rohrmeier, M. 2009. ‘Tacit tonality: Implicit learning of context-free harmonic structure.’ Paper presented at ESCOM, 12-16 August 2009, University of Jyväskylä, Finland.

Cross, I., & Woodruff, G. E. 2008. ‘Music as a Communicative Medium.’ In R. Botha and C. Knight (eds.), *The Prehistory of Language*, pp. 113-144. Oxford: Oxford University Press.

Davies, S. 2003. *Themes in the Philosophy of Music*. New York: Oxford University Press.

Dennett, D. 1991. *Consciousness Explained*. Boston: Little Brown.

Deutsch, D. (ed.) 1992. *The Psychology of Music*, second edition. San Diego: Academic Press.

Deutsch, D. 1982. 'Organizational Processes in Music.' In M. Clynes (ed.), *Music, Mind, and Brain: The Neuropsychology of Music*, pp. 119-136. New York: Plenum Press.

Deutsch, D. 1980. Music Perception. *The Musical Quarterley*, 66 (2): 165-179.

Devitt, M., & Sterelny, K. 1987. *Language & Reality: An Introduction to the Philosophy of Language*. Oxford: Basil Blackwell.

Dissanayake, E. 2008. If music is the food of love, what about survival and reproductive success? *Musicae Scientiae* Special Issue 2008: 169-195.

Dodd, J. 2000. Musical works as eternal objects. *The British Journal of Aesthetics*, 40 (4): 424-440.

Duckworth, W. 1999. *20/20: 20 New Sounds of the 20th Century*. New York: Schirmer Books.

Erpf, H. 1967. *Form und Struktur in der Musik*. Mainz: B. Schott.

Fink, R. 2001. 'Going Flat: Post-hierarchical Music Theory and the Musical Surface.' In N. Cook and M. Everist, *Rethinking Music*, pp. 102-137. New York: Oxford University Press.

Fodor, J. 1998. The Trouble with Psychological Darwinism. *London Review of Books*, 20 (2): 11-13.

Grice, H. P. 1991. *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.

Gillon, B. S. 2008. On the Semantics/Pragmatics Distinction. *Synthese*, 165 (3): 373-384.

Hannon, E. E., and Johnson, S. P. 2005. Infants use meter to characterise rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, 50: 354-377.

Harwood, D. L. 1976. Universals in Music: A Perspective from Cognitive Psychology. *Ethnomusicology*, 20 (3): 521-533.

Hauser, M. D., Chomsky, N., and Fitch, W. T. 2002. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298:1569-1579.

Huron, D. 2003. Is Music an Evolutionary Adaptation? In I. Peretz and R. Zatorre, *The Cognitive Neuroscience of Music*, pp. 57-76. Oxford: Oxford University Press.

Ingarden, R. 1986. *The Work of Music and the Problem of Its Identity*, trans. A. Czerniawski, ed. J. G. Harrell. Houndmills: Macmillan Press.

Jackendoff, R. 2009. Parallels and Nonparallels between Language and Music. *Music Perception*, 26 (3): 195-204.

Jackson, F. 1982. Epiphenomenal Qualia. *Philosophical Quarterly*, 32: 127-136.

Juslin, P. N., and Laukka, P. 2003. Communication of emotions in vocal expression and musical performance: Different channels, same code? *Psychological Bulletin*, 129: 770-814.

Juslin, P. N., and Sloboda, J. A. (eds.) 2001a. *Emotion and Music: Theory and Research.* Oxford: Oxford University Press.

Juslin, P. N., and Sloboda, J. A. 2001b. 'Music and Emotion: Introduction.' In P. N. Juslin and J. A. Sloboda (eds.), *Emotion and Music: Theory and Research*, pp. 3-20. Oxford: Oxford University Press.

Kehler, A. 2002. *Coherence, Reference, and the Theory of Grammar.* Stanford: CLSI Publications.

Kivy, P. 2002. *Introduction to a Philosophy of Music.* New York: Oxford University Press.

Kivy, P. 1990. *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca and London: Cornell University Press.

Knösche, T. R., C. Neuhaus, J. Haueisen, K. Alter, B. Maess, O. Witte, and A. Friederici. 2005. Perception of phrase structure in music. *Human Brain Mapping*, 24 (4): 259-273.

Koopman, C., and Davies, S. 2001. Musical meaning in a broader perspective. *The Journal of Aesthetics and Art Criticism*, 59 (3): 261-273.

Kramer, L. 2002. *Musical Meaning: Toward a Critical History*. Berkeley: University of California Press.

Lerdahl, F. 2009. Genesis and Architecture of the GTTM Project. *Music Perception*, 26 (3): 187-194.

Lerdahl, F. 2001. *Tonal Pitch Space*. New York: Oxford University Press.

Lerdahl, F. 1988. 'Cognitive Constraints on Compositional Systems.' In J. Sloboda (ed.), *Generative Processes in Music*, pp. 231-259. Oxford: Clarendon Press.

Lerdahl, F., and R. Jackendoff. 2006. The capacity for music: What is it, and what's special about it? *Cognition*, 100: 33-72.

Lerdahl, F., and R. Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.

Lerdahl, F., and R. Jackendoff. 1981. On the Theory of Grouping and Meter. *The Musical Quarterly*, 67 (4): 479-506.

Lerdahl, F., and C. Krumhansl. 2007. Modeling Tonal Tension. *Music Perception*, 24 (4): 329-366.

Levitin, D. 2006. *This is Your Brain on Music: the Science of a Human Obsession.* New York: Dutton.

Levinson, J. 1990. *Music, Art and Metaphysics.* Ithaca: Cornell University Press.

Lewis-Williams, D. 2002. *The Mind in the Cave: Consciousness and the Origins of Art*. London: Thames & Hudson.

List, G. 1971. On the Non-Universality of Musical Perspectives. *Ethnomusicology*, 15 (3): 399-402.

Lindsey, G. 1994. Review of Raffman (1993). *Journal of Linguistics*, 30 (2): 575-577.

McClary, S. 1991. *Feminine Endings: Music, Gender, and Sexuality.* Minneapolis: University of Minnesota Press.

Meyer, L. B. 1960. Universalism and Relativism in the Study of Ethnic Music. *Ethnomusicology*, 4 (2): 55-59.

Meyer, L. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Middleton, R. 1990. *Studying Popular Music*. Philadelphia: Open University Press.

Miller, G. 2000. 'The Evolution of Music Through Sexual Selection.' In N. Wallin, B. Merker, and S. Brown (eds.), *The Origins of Music*, pp. 315-328. Cambridge, MA: MIT Press.

Mithen, S. 2005. *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. London: Weidenfeld & Nicolson.

Mohanan, K. P. 1995. 'The Organization of the Grammar.' In J. A. Goldsmith (ed.), *The Handbook of Phonological Theory*, pp. 24-69. Cambridge, MA: Blackwell.

Nattiez, J. J. 2004. *The Battle of Chronos and Orpheus: Essays in Applied Musical Semiology*, trans. J. Dunsby. Oxford: Oxford University Press.

Nattiez, J. J. 1990. *Music and Discourse: Toward a Semiology of Music*.

Nettl, B. 2005 [1983]. *The Study of Ethnomusicology: Thirty-one Issues and Concepts.* Urbana: University of Illinois Press.

Omerzel-Terlep, M. 2000. 'Slovenia.' In T. Rice, J. Porter, and C. Goertzen (eds.), *The Garland Encyclopedia of World Music*, vol. 8, pp. 911-914. New York: Garland.

Patel, A. D. 2008a. *Music, Language, and the Brain*. Oxford: Oxford University Press.

Patel, A. D. 2008b. Talk of the tone. *Nature*, 453: 726-727.

Patel, A. D. 2006. Musical Rhythm, Linguistic Rhythm, and Human Evolution. *Music Perception*, 24 (1): 99-103.

Patel, A. D., Peretz, I., Tramo, M., and Labrecque, R. 1998. Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language*, 61 (1): 123-144.

Peretz, I. 2006. The nature of music from a biological perspective. *Cognition*, 100: 1-32.

Peretz, I., and Coltheart, M. 2003. Modularity of music processing. *Nature Neuroscience*, 6 (7): 688-691.

Phillips-Silver, J., and Trainor, L. J. 2005. Feeling the beat: Movement influences infants' rhythm perception. *Science*, 308 (5727): 1430.

Pinker, S. 1997. *How the Mind Works*. New York: Norton.

Pinker, S. 1994. *The Language Instinct*. New York: Morrow.

Polger, T. W., and Flanagan, O. 1999. 'Natural Answers to Natural Questions.' In V. G. Hardcastle (ed.), *Where Biology Meets Psychology: Philosophical Essays*, pp. 221-247. Cambridge, MA: MIT Press.

Premack, D., and Woodruff, G. 1978. Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences*, 1 (4): 515-526.

Raffman, D. 1993. *Language, Music, and Mind*. Cambridge, MA: MIT Press.

Raffman, D. 1988. Toward a Cognitive Theory of Musical Ineffability. *The Review of Metaphysics*, 41 (4): 685-706.

Rothstein, W. 1989. *Phrase Rhythm in Tonal Music*. New York: Schirmer Books.

Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson's University Library.

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

Scruton, R. 1987. Analytical philosophy and the meaning of music. *The Journal of Aesthetics and Art Criticism*, 46: 169-176.

Shannon, C., and Weaver, W. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Shepard, R. 2001. 'Cognitive Psychology and Music.' In P. R. Cook (ed.), *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*, pp. 21-35. Cambridge, MA: MIT Press.

Shepard, R. 1964. Circularity in Judgments of Relative Pitch. *The Journal of the Acoustical Society of America*, 36 (12): 2346-2353.

Stalnaker, R. C. 1970. Pragmatics. *Synthese*, 22 (1/2): 272-289.

Swain, J. 1997. *Musical Languages*. New York: W. W. Norton.

Tagg, P. 1979. *Kojak: 50 Seconds of Television Music*. Göteborg: Skrifter från Musikvetenskapliga institutionen, 2.

Tarasti, E. (ed.) 1995. *Musical Signification: Essays in the Semiotic Theory and Analysis of Music.* Berlin: Mouton de Gruyter.

Tarski, A. 1944. The Semantic Conception of Truth, and the Foundations of Semantics. *Philosophy and Phenomenological Research*, 4 (3): 341-376.

Temperley, D. 2000. Meter and Grouping in African Music: A View from Music Theory. *Ethnomusicology*, 44 (1): 65-96.

Tenney, J., and Polansky, L. 1980. Temporal Gestalt Perception in Music. *Journal of Music Theory*, 24 (2): 205-241.

Trainor, L. J., and Trehub, S. E. 1992. A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, 18: 394-402.

Treitler, L. 1993. 'Gender and Other Dualities of Music History.' In R. Solie (ed.), *Musicology and Difference: Gender and Sexuality in Music Scholarship*, pp. 23-45. Berkeley: University of California Press.

Trehub, S. E. 2003. 'Musical Predispositions in Infancy: an Update.' In I. Peretz and R. Zatorre (eds.), *The Cognitive Neuroscience of Music*, pp. 3-20. Oxford, Oxford University Press.

Trehub, S. E., and Hannon, E. E. 2006. Infant music perception: Domain-general or domain specific mechanisms? *Cognition*, 100: 73-79.

Williams, A. 2001. *Constructing Musicology*. Aldershot: Ashgate.

Wertheimer, M. 1923. *Drei Abhandlungen zur Gestalttheorie*. Erlangen: Philosophischen Akademie.

Wittgenstein, L. 1961 [1922]. *Tractatus Logico-Philosophicus*, trans. D. F. Pears and B. F. McGuiness. London: Routledge & Kegan Paul.