Nikolaus Steinbeis: Investigating the Meaning of Music using EEG and FMRI. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences, 2008 (MPI Series in Human Cognitive and Brain Sciences; 100)

Investigating the meaning of music using EEG and fMRI

Der Fakultät für
Biowissenschaften, Pharmazie und Psychologie
der Universität Leipzig
eingereichte

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
Dr. rer. nat.

vorgelegt
von Nikolaus Steinbeis
geboren am 16. April 1979 in München

Dekanin:
Prof. Dr. Annette Beck-Sickinger
Gutachter:
Prof. Dr. Erich Schröger
PD Dr. Stefan Koelsch
Dr. Mireille Besson

Tag der Verteidigung: 9. März 2008

Acknowledgements

At this point I would like to take the opportunity to thank all those who have contributed in their own way to the inception, execution and completion of my dissertation.

First of all thanks goes to my supervisor, PD Dr. Stefan Koelsch for his knowledgeable supervision and the enormous scientific freedom granted to develop the ideas contained in this dissertation. I would also like to thank Prof. Erich Schröger for kindly offering to act as reviewer. Thanks also goes to Dr. Mireille Besson for reviewing this work. She has been a source of inspiration and figure of guidance in keeping an open mind when looking at data and I have enjoyed the many fruitfuls discussions.

A special thanks goes to all the members of the music group, especially Daniela Sammler and Clemens Maidhof, both of whom have thoroughly scrutinised this work and who apart from everything are treasured roommates and friends. I am also grateful to Sonja Kotz for her unceasing willingness to discuss data and beyond.

I thank Sylvia Stasch for all the help with the EEG data acquisition and the many interesting and entertaining chats we had during the course of it. Thanks also go to Heike Boethel, Mandy Naumann, Simone Wipper and Petra Erz for help with the EEG and the fMRI data. Kerstin Flake has also been hugely helpful (and patient) with many of the images contained in this dissertation.

A final thanks goes to my family and friends who have been the most important contributors, not with knowledge but with love and support.

Contents

Preface

“*La musique exprime ce qui ne peut pas être dit*
et sur quoi il est impossible de rester silencieux.”
Victor Hugo (1802-1885)

Both vocal and instrumental music are used in a vast variety of different settings, functioning as accompaniment to rituals and ceremonies, cementing social interactions [1] or quite simply providing a source of immense pleasure to listeners and performers alike. Its similarities to language, in that it can be vocally produced, is highly complex and infinitely variable, subject to a particular set of implicitly learned rules and the spontaneous use of which is presumably exclusively human, have tempted towards making several assumptions with regards to its evolutionary purpose and potential function. One of these assumptions has centred around the question of what musical signals can communicate. Is music capable of communication? If so, what is the nature of musical communication? Does music possess meaning? Is this meaning comparable to the kind of semantic meaning inherent to language?

Apart from agreeing on the notion that, like language, music is a particular instance of a *sign system*, ideas diverge on how these signs are understood. For instance, several scholars have argued that music conveys meaning, which can be unanimously perceived (Meyer, 1956; Sloboda, 1986; Swain, 1997). As a result, numerous possibilities of the type of meaning that music may convey have been entertained, such as musical mood suggestion, sounds mimicking objects, associations with the music, or the interplay of structural features. In spite of these sometimes very explicit formulations of the mechanisms by which each type of meaning may be conveyed (e.g. Meyer, 1956), the empirical evidence to date has been surpri-

[1] i.e. Fostering group cohesion and mother-child bonding (Wallin, Merker, & Brown, 2000)

singly scant. The present thesis is an attempt to bridge the gap between theory and evidence, by virtue of investigating the basis for several of these hypothesised routes to meaning in music.

In spite of an initial attempt showing the general ability of music to convey meaningful information (Koelsch et al., 2004), there is still a lack of thorough research into the actual mechanisms hypothesised to underlie musical meaning. This dissertation relies on biological theories of auditory communication, musicological notions of musical communication and the experimental literature on the cognitive (neuro)science of music, to draw up a systematic research agenda on some of the mechanisms underlying musical meaning.

Two of the mechanisms by means of which meaning has been proposed to arise will be focussed on, specifically the relationship between meaning and emotion and the role of musical structure. An additional issue of concern to the present thesis was whether the meaning conveyed by music is represented in a manner similar to meaning conveyed by language, which was dealt with by means of comparing neural responses to both music and language targets in an affective priming paradigm. Thus, the thesis covers three separate empirical issues related to meaning in music and was structured accordingly into a first part adressing the relationship of meaning and emotion (Experiment 1-3), a second part adressing the link between musical structure and meaning (Experiment 4), and a third part comparing the neural representations of meaning in music and language (Experiment 5 & 6). A brief summary of each chapter will ensue, followed by an outline of the experiments and the main findings.

Chapter 1 gives an introduction to the basis of auditory communication of meaning. A working definition of communication will be given and after outlining the role of vocalizations as communications of meaning in the animal kingdom, the discussion turns to verbal and non-verbal communication in humans as exemplified by language, prosody and music. It is argued that referential meaning is a property shared by other species and something that music may also be capable of.

Because language has been the primary focus for the investigation of meaning, Chapter 2 gives a detailed account of semantic processing. After introducing some models of word recognition, the empirical paradigms and the type of neuroscientific evidence used to indicate semantic processing are discussed. This provides the reference point for discussing meaning in music.

Chapter 3 gives a detailed account of the processes underlying the perception and cognition of music and introduces a recently proposed model of music processing. This is followed by a discussion of what meaning in music can and cannot entail, leading onto a further cagtegorisation of meaning into referential and non-referential meaning. In addition to a thorough overview of how meaning in music has been argued to arise, a review of event-related potentials in response to music is given, serving as background with which data from the following empirical part is discussed.

Chapter 4 compares language and music with regards to similiarites in the cognitive operations involved in their processing. Recent proposals include a framework for resource sharing between the two domains (Patel, 2003). This framework is introduced and some evidence in favour is discussed with view to include other cognitive operations than the ones hitherto considered.

The present thesis also investigated the role of musical expertise in processing meaning in music. Chapter 5 provides an outline of the mechanisms underlying neuronal plasticity and presents both anatomical and functional evidence for neural changes resulting from musical training. Behavioural differences between the two groups are also discussed, suggesting a much smaller discrepancy than the neural data implies.

After describing the methods used in the dissertation in Chapter 6 and outlining the research questions in Chapter 7, Chapter 8 adresses the relationship of meaning and emotion in music. By means of an affective priming paradigm, single chords varying either in harmonic roughness (Experiment 1), intervals (Experiment 2), or instrumental timbre (Experiment 3), which are therefore either more or less pleasant, are used to prime the processing of subsequent emotional target words. It is shown that each of the three musical parameters, varying in expressed emotions, can significantly influence the processing of emotional word targets, regardless of musical training. The ERP data suggest that this is due to the affective meaning communicated by the single chords, providing the first piece of evidence for a link between meaning and emotion in music.

Chapter 9 adresses the long-standing hypothesis that tension-resolution patterns are meaningful to listeners familiar with Western music. By means of a paradigm using harmonic chord sequences with an expected and an unexpected final progression, as well as sentences with varying types of violations (e.g. syntactic, semantic, correct), it is shown that simultaneous presentation of semantic violations in language diminishes a music-related potential. It is assumed that music and language share processing resources on several levels and this speaks for shared resources on a semantic level as well as for a role of tension-resolution patterns in musical meaning.

The final empirical issue is adressed in Chapter 10 and looks at the comparibility of neural representations for meaning in music and language. The results indicate that under the appropriate circumstances, musical targets can elicit ERPs highly reminiscient of those elicited by semantic violations in language, but only for musically trained subjects (Experiment 5). A subsequent fMRI study (Experiment 6) was designed with the aim to uncover the neural correlates of processing meaning in music as well as to dissociate two competing explanations for the data of Experiment 5 (response conflict vs semantic conflict). In conjunction with the ERP data, results from the fMRI suggest some degree of comparibility in the neural representations of meaning in music and language, but also important functional differences.

Chapter 11 discusses the findings reported in the preceding three chapters with respect to the literature on music, meaning and emotional processing.

Part I

Theoretical part

Chapter 1

The nature of communication

1.1 Bases of auditory communication

Communication is undeniably a fundamental aspect of life. Whether inadvertent or intentional, all organisms send and receive signals by means of which they give information that can be used by their living environment to modify behaviour. There are obvious evolutionary advantages to effective communication, presumably a reason why the tools for communication, both for production and comprehension, have developed into highly sophisticated and complex systems of which human language is arguably the pinnacle. Music has also been argued to function as a non-verbal means of communication (Sloboda, 1986; Swain, 1997). Whereas this is debatable from an evolutionary perspective[2], what musical communication involves is still an under-researched area and the topic of the present thesis.

This chapter will give a definition of communication and set out how this term will be used throughout the present work. Subsequently a brief outline of communication in the animal kingdom will be given, which, because of the present research focus on language and music, will concentrate on auditory communication. This will be followed by discussing communication in humans, both verbal such as language, and non-verbal such as prosody and music.

[2]In the face of the communicative efficency of language it is unlikely that music evolved for explicit communicative purposes.

1.1.1 What is communication

Communication has been defined with subtle variations depending on which scientific discipline has used the term (Hauser, 1996). The present defintion postulates that communication consists of a signal which is sent by one and received by another organism. This signal contains information, which is understood as such by the receiver and which usually leads to some modification in behaviour. Throughout the subsequent review the issue of what kind of information is contained in the signal (e.g. vocalizations, speech, prosody and music) and how this is used to communicate will be adressed.

1.2 Vocalizations in the animal kingdom

Whereas during animal communication every availabe sense is likely to be exploited, the current interest in language and music guides our focus towards auditory communication. There are a multitude of these in the animal kingdom such as birdcalls, whalesong, or alarm calls in chickens (for a review, see Wallin et al., 2000), but the present section will focus on vocalizations in non-human primates, because relatively compelling evidence for the use of referential meaning in vocalizations has amassed. *Referentiality* in communication is used in the sense that vocalizations appear to map onto salient objects and events in the environment.

Observations in the field and systematic empirical studies (Seyfarth, Cheney, & Marler, 1980) have shown that vervet monkey vocalizations do not just reflect a signaller's affective state, but are also capable of referring to specific types of predators (e.g. big cats, birds of prey and snakes). Playback experiments of recorded predator warning calls were capable of eliciting appropriate behavioural responses in conspecifics in the clear absence of an actual predator, showing a highly developed and interactive system of communication and related adaptive behaviour.

Because these early findings have suggested such a high degree of sophistication in vocalizations, theoretical frameworks have become increasingly refined with regards to the necessary and sufficient conditions to allow for successful communication to take place (Seyfarth & Cheney, 2003b). Currently it is argued that natural selection has favoured callers who vocalize to affect the behaviour of listeners who in turn extract information from the vocalizations to represent the environment

(Seyfarth & Cheney, 2003b). This has been shaped by evolution, particularly where there is a significant overlap in the evolutionary interests of signaler and recipient (for a review on predator calls in vervet monkeys, see Seyfarth & Cheney, 2003a).

There are important similarities and differences between the kind of referential communication used by non-human primates and human language. The obvious similarity starts at the mere use of referential sounds. Claims to the contrary, whereby animal vocalizations are merely affective utterances and do not posess the inherent referential properties of human language (i.e. establishing the relation between words and the objects or events they represent) have been called into question (Seyfarth & Cheney, 2003b). Both, vocalizations and language reliably refer to an external event, regardless of the underlying mechanisms. Thus, animal calls, even though much less sophisticated and exhaustively so, share a basic capacity for referentiality with human language.

A major difference in the use of communication between humans and non-human primates is the role of goal understanding and intentionality (i.e. whether the sender of a communicative signal is aware of its perception in the receiver). Influencing others' behaviour by changing what they know, think, believe, or desire has been proposed as a primary function of language (Pinker, 1994). This presupposes the speaker's awareness of the listener's mental contents, also known as possessing a theory of mind (TOM; Premack & Woodruff, 1978). Recent evidence suggests that, even though animal calls may alter behaviour, this occurs inadvertantly, as the signaller is unaware of its effects (for a review, see Seyfarth & Cheney, 2003a). In spite of this, some tentative evidence has suggested the presence of precursors of a theory of mind in chimpanzees, such as following eye gaze (Tomasello, Call, & Hare, 1998) and possession of knowledge of what conspecifics may or may not have seen (Hare, Call, & Tomasello, 2001). Most recently it has been shown that orang-utangs modify their gestures according to their perception of their audience's comprehension (Cartmill & Byrne, 2007), suggesting a more sophisticated capacity of representing others' intentions in non-human primates than previously assumed.

Thus, it appears as if there are instances of vocal communication in non-human primates whereby highly specific information is being transmitted and understood as such, establishing a system of referential meaning, possibly mediated by affective states. There are obvious limits to the breadth and diversity of these signals, which may be due to limitations of more general cognitive skills, such as memory.

However, the additional fact that communication in animals appears to occur without any understanding of the receiver's state of mind, appears to be a fundamental difference to the way that humans use language.

1.3 Communication in humans: Language, prosody, and music

1.3.1 Language

The most pervasive and efficient means of communication for humans is language. It is understood as a culturally specific communication system enabling the communication of simple factual statements, such as about the weather or the colour of a flower, inner mental states, such as emotions, beliefs and desires, as well as to entertain hypothetical scenarios, such as: "What if Margaret Thatcher came back into power?". Most strikingly all languages seem to be able to do so in an infinite variety of ways, using words of the same meaning interchangeably and in varying order. Even though, as argued above, animals do use effective communicative systems, these appear to lack the rich expressive and open-ended capacity of human language. In terms of communicative systems, this has been proposed to be a primary feature differentiating language from other forms of communication in the animal kingdom (Hauser, Chomsky, & Fitch, 2002; Fitch & Hauser, 2004).

As the specific components of language will be discussed in greater detail in the following chapter, the present summary will confine itself to a very brief outline of some of the defining characteristics of language as recently put forward by Hauser et al. (2002), with an emphasis on its inherently referential nature.

A recent proposal on the evolution of the language faculty (Hauser et al., 2002) has suggested that in a broad sense, what defines language is the presence of (1) a sensory-motor system, most commonly auditory-articulatory, (2) a conceptual-intentional system, and (3) the computational mechanisms for recursion, the last of which provides the capacity for an infinite number of expressions derived from a limited set of elements. It is argued by Hauser et al. (2002, p. 1571) that the computational system of recursion "...generates internal representations and maps them into a sensory-motor interface by the phonological system, and into the conceptual-intentional interface by the semantic system". The degree to which any of these three aspects may or may not be shared by other animals is still under considerable debate (Fitch & Hauser, 2004; Gentner, Fenn, Margoliash, & Nusbaum, 2006).

The processing of an expression generated by the sensory-motor system, and more importantly by the conceptual-intentional system, brings one of the core aspects of language to the fore, namely that each expression entails the mapping of sound to meaning. This brings the discussion to the aspect of most importance to the present work, namely the semantic system. This system has been argued to consist of all the different types of information encountered during one's life, ranging from direct or indirect experience through instruction to its mere verbal transmission (Kutas & Federmeier, 2000). Once this information is stored, it constitutes knowledge and is retained in memory. Semantic memory[3] in particular represents the system from which information is retrieved whenever, in the case of language, a linguistic cue in the form of a spoken, written, or signed word is processed. Semantic memory is enormously complex and flexible. Even though in principle it might be comparable to the use of the vocalizations described above, human language far exceeds the number of possible combinations that meaning and sound can have.

Thus human language comprises a set of basic underlying computational mechanisms, one of which entails the mapping of sounds to their meaning. Even though the ultimate production and real-life comprehension of language is much more complex than simple sound-to-meaning mapping, this semantic or referential system (whereby a sign refers to a meaning) is of most interest to the present work. Humans have an elaborate system by means of which the incoming perceptual information is automatically mapped onto its meaning, which allows for enormously effective and versatile communication. The subsequent section will discuss a sub-component of auditory language comprehension and production, namely prosody and will then venture onto discuss the uniquely human cultural artifact of music.

[3] The concept of memory has been subdivided into a so-called semantic memory systems, which includes factual knowledge about the world and an episodic memory system, which includes all events related to oneself (Tulving & Donaldson, 1972). Like all distinctions, the boundaries between the two systems are necessarily fluid.

1.3.2 Prosody

Spoken language is different to written language in several important ways. For instance the onset and offset of words in a sentence is rarely marked by an acoustic break, but rather occurs continuously in a single stream of speech. More importantly however, spoken language is subjected to a pattern of stresses, intonation, speed and energy. This is also known as prosody and has been shown to contain a significant amount of information with regards to the syntactic structure of a sentence (Friederici & Alter, 2004), as well as the emotional state of the speaker (Schirmer & Kotz, 2006).

Studies have shown that prosodic information can directly influence the syntactic parser, whereby prosodic phrase boundaries will either lead to further syntactic integration into the current phrase or not. When prosodic cues go against the initial parsing preference, this leads to significant processing costs, presumably at the level of lexical re-access and subsequent structural revision (Steinhauer, Alter, & Friederici, 1999). Thus one function of prosody appears to be to further structure the auditory speech input, which is then used for speech comprehension.

An additional aspect of prosody that has come under increased scientific focus recently is the communication of emotion. It has been assumed that vocalizations are closely linked to physical parameters associated with emotional states (e.g. muscle tension, respiration rate, blood pressure; Scherer, 1989). Thus vocalizations and particularly vocal expressions in humans, including prosody, have been argued to directly reflect emotions (Scherer, 1989). Emotion has been suggested to be communicated primarily through four acoustic parameters: 1) vocal intensity (loudness); 2) vocal frequency (F0 and its harmonics); 3) vocal quality, primarily affected by phonatory and articulatory settings, as well as 4) vocal resonance. Systematic study of the production and perception of emotional prosody has yielded a consistent picture, whereby 14 different emotions vocalized by professional actors could be reliably classified and differentiated along acoustic parameters (Banse & Scherer, 1996). These initial findings have recently been confirmed and extended across a range of different cultures and languages, indicating that some of the identified acoustic parameters may be of universal emotional significance (Scherer, Banse, & Wallbott, 2001).

Another line of research on the perception of emotional prosody has focussed on more global aspects, such as the influence of affective speech on semantic processing (Schirmer, Kotz, & Friederici, 2002, 2005; Schirmer & Kotz, 2003, 2006; Schirmer, Zysset, Kotz, & Cramon, 2004). This series of studies could demonstrate that varying the emotional prosodic expression of affectively neutral sentences, could prime the subsequent processing of emotionally valenced words (Schirmer et al., 2002). Thus affective prosody could be shown to influence semantic processing, providing evidence in support of the notion that prosodic cues are directly relevant to the understanding of meaning expressed in speech, representing an additional level of referential, albeit non-verbal information contained in the speech signal.

1.3.3 Music

The previous review of how meaning can be derived from a variety of acoustic cues, a communicative mechanism already present in the animal kingdom, and at its most refined in language, including both verbal and non-verbal cues, leads to the present discussion of music, a complex and arguably exclusively human phenomenon, which has frequently been looked at as a non-verbal means of communication (Meyer, 1956; Sloboda, 1986; Swain, 1997; Patel, 2003; Koelsch & Siebel, 2005; Koelsch, 2005).

In basic descriptive terms, music is an art form, which consists of a series of sounds and silences expressed in time. The sounds are organized in terms of pitch (including melody and harmony), rhythm (including tempo and meter), and other qualities such as the timbre, dynamics, instrumentation and variables that can be influenced by the performer. The kind of music and the contexts in which music can occur differ enormously between cultures (Cook, 1998).

How and for what purpose music may have evolved has been of increased scientific interest over recent years (Wallin et al., 2000; Fitch, 2005, 2006; Hauser & McDermott, 2003; McDermott & Hauser, 2005). Questions, such as whether spontaneous and unprompted music production actually occurs anywhere in the animal kingdom, whether innate musical preferences exist or whether music was an evolutionary precursor to language or the other way round, are all topics of intense scholarly debate, which will not be discussed further here. Rather this section aims to review the idea that music can be seen as a means of communication. This issue in particular has attracted philosophers and musicologists alike (for discussions on the topic, see Jones & Holleran, 1991; Juslin & Sloboda, 2003), and there appear to be two aspects to this, which will be briefly discussed in turn. The first involves the notion of communication through the composer or the performer, whereas the second deals with the notion of inherently communicative musical properties. Questions arising out of the latter issue will provide the basis for the empirical research presented in this thesis.

It has been argued in the past that music might be viewed as part of a communication system in which the composers code musical ideas in notation, performers recode from the notation to a musical signal, and listeners recode from the acoustic signal to ideas (Juslin, 2003; Kendall & Carterette, 1990). In Western tradition, music at some stage (either live or recorded) requires a performer to be heard. Performers will either play music that has been composed previously or may improvise. Either way, unlike visual arts, the musical expression will vary between each performance and the performer will typically have some intention of what he or she would like to communicate to the listener (Sloboda & Lehmann, 2001; Juslin, 2003; Palmer, 1989, 1997). These communications can entail emphasising structural aspects of the music, reflecting the performer's conceptual grasp of the piece (Clarke, 1988; Palmer, 1989, 1997), or rather involve giving the music *expressive shape* (Repp, 1998). Empirical work has focussed predominantly on providing answers to two questions: (1) the accuracy of communication; and (2) the communicative code used by performers and listeners (Juslin, 2003). Most importantly however it appears as if there is a consistent overlap between the performer's communicative intentions and its being understood as such by the listener (for a review, see Juslin, 2003).

That musical meaning can arise simply out of musical propoerties has also been proposed and can be divided further into referential and non-referential communication. Referential meaning in music has been characterized both in terms of reference to an extra-musical event, as well as in terms of reference to another section or theme within the same musical piece (Sloboda, 1986). A most rigorously descriptive account of what emotions music can refer to was set out by D. Cooke (1959), which postulates, with moderate success that the intervals of the diatonic scale suggest different emotional qualities (e.g. minor third - sadness; major third - happiness).

Another conceptualisation of meaning in music has been posited based on its structure and lack of external reference or non-referentiality. The seminal work of L.B. Meyer (1956) argues that musical structure is capable of communicating both meaning and emotion, by means of musical tension-resolution patterns. These patterns reflect the listeners's expectations and the manipulation of which gives rise to both affective responses as well as an understanding of music.

The evidence on both referential and non-referential communication in music is unnervingly scant. To date, there is no rigorous empirical test of the mechanisms by means of which music can communicate to the listener, in spite of evident hypothetical links. The present thesis is an attempt to fill this void by means of exploring the variety of hypothesised routes to meaning in music.

1.4 Summary

The present overview provided a rough sketch of instances of communication in animals and humans. Animals are capable of referential communication, albeit without the necessarily intent to do so or any notion that conspecifics understand. Humans in turn have a hugely complex and diverse system for communicating an infinite variety of things by means of language. Prosody is an additional aspect of language, serving syntactic structure comprehension, as well as communicating emotions. Music has frequently been viewed as a non-verbal means of communication, including the composer's or performer's intentions, as well as the ability to refer to external events or internal musical structure. This thesis adresses the idea that music can communicate by means of several routes. After reviewing the relevant literature on the communication of meaning in language, the hypotheses for several routes to meaning in music will be explored.

Chapter 2

Language processing

As has been argued in the previous chapter, a key feature of language is the operation of a conceptual-intentional system subserved by semantic memory. Because it is of central importance to a complete understanding of language function and because semantic memory is a crucial aspect of human cognition in general, this topic has received substantial scientific attention. The present chapter will provide a brief overview of relevant theoretical models and empirical data on semantic processing. Studies investigating the neural correlates which underlie the comprehension of semantic aspects of language will also be reviewed. This will serve as crucial conceptual basis for considering meaning in music.

2.1 Theoretical Overview

Human language is an enormously complex phenomenon, both for the speaker and the listener. When trying to make sense of the acoustic speech signal, an enormous variety of different information has to be processed, including segmental phonemes, suprasegmental phonological information, as well as syntactic structure and semantic meaning. The present section will focus on a circumscribed aspect of language comprehension, namely word recognition, which will provide an outline of the processes underlying semantic aspects of language comprehension.

2.1.1 Word recognition

Words represent the most fundamental unit of semantic comprehension. To understand how semantic processing in language works it makes most sense to begin by reviewing some theoretical perspectives on word recognition.

Word recognition has been argued to comprise three basic functions: access, selection, and integration (Marslen-Wilson, 1987; Zwitserlood, 1994, 1998). Lexical access concerns the relationship between sensory-perceptual input and the mental lexicon, where word meanings are stored. If the sensory input is sufficient for word identification the word with all its associated properties (e.g. morphology, word class, meaning) will get selected from the lexicon. This is known as lexical selection and essentially involves singling out the element that best matches the given perceptual input, from all the ones activated. This step is of particular importance for word disambiguation and necessarily implies drawing on further contextual knowledge provided (Hagoort & Brown, 2000). The final stage is known as lexical integration, which involves the integration of the retrieved and identified word into the existing linguistic context, binding syntactic and semantic information into the whole utterance.

Wheras the theoretical consensus with regards to the operation of these stages is relatively unanimous, the theories diverge on the issue of how the stages relate to one another and their degree of modularity (Zwitserlood, 1998). Both modular/parallel (Marslen-Wilson, 1984, 1987) and interactive accounts (McClelland & Rumelhart, 1981; McClelland & Elman, 1986; Rumelhart & McClelland, 1982) have been formulated, especially on the topic of word recogntion. In the so-called COHORT model (Marslen-Wilson, 1984) it is proposed that lexical access entails the activation of all word forms which may be compatible with the initial sensory input of a spoken word. The more information becomes availabe, the more rival candidates are eliminated. Once the information is sufficient, the candidate will be retrieved from the mental lexicon and integrated into the sentence. Whereas earlier versions of the model allowed for semantic and syntactic constraints on lexical access and retrieval (Marslen-Wilson & Tyler, 1980), more recent versions have argued that syntactic and semantic information of multiple elements contained in the speech stream are assessed in parallel by the process responsible for integration. Even though there has been some evidence in favour of this modular approach

(Zwitserlood, 1989), there has been some theoretical opposition favouring an interactive language approach (McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). This latter approach predicts that all different sources of information interact with each other during word processing. In the so-called TRACE model, word elements become activated with both lower (e.g. phonemes) and higher levels of representation. That way, contextual information can affect the activation level of words at any point, thereby incrementally increasing the likelihood that a contextually appropriate word will be selected.

A recent neurocognitive model of language comprehension has tried to integrate the rivalling accounts (Friederici, 2002). By relying predominantly on electrophysiological and functional imaging data, an account is proposed, whereby autonomous, parallel processes occur at early stages and interactive processes at later stages of comprehension.

2.2 Empirical findings

2.2.1 Semantic priming

The primary method of investigating the processing of meaning, both at the single word as well as at the sentence level is semantic priming. The term *priming* refers to an improvement in performance in a perceptual or cognitive task, measured relative to an appropriate baseline, subsequent to prior experience or familiarity with the context.

Semantic priming is the improvement in speed and/or accuracy to respond to a stimulus (typically a word or an image), when preceded by a semantically related stimulus (e.g. bread-butter), compared to when preceded by a semantically unrelated stimulus (e.g. horse-butter). The term semantic priming typically refers to priming caused by a mixture of genuine semantic, but also associative relations (e.g. cow-horse). Whereas the stimulus to which the response has to be made is known as the *target*, the preceding stimulus is referred to as the *prime*. Behavioural tasks typically measuring improved performance include the *lexical-decision task*, whereby target words have to be classified into genuine words or so-called *pseudo-*

words, consisting of a random string of letters, as well as other types of decisions to be made on the nature of the target.[4]

There are two prominent models attempting to account for the findings on semantic priming (McNamara, 2005), spreading models and multistage activation models. Spreading models have provided some of the earliest accounts of semantic priming (Quillian, 1967; Collins & Loftus, 1975) and in spite of their variety have agreed on the following basic theoretical tenets: (a) the retrieval of an item from memory activates its internal representation; (b) activation of such a representation will spread to related representations; and (c) the remaining activation of related representations will facilitate their subsequent retrieval. Multistage activation models on the other hand argue for (a) the existence of multiple levels of lexical-semantic representations (e.g. visual features, letters, words, semantic representations); (b) the operation of excitatory feedforward and feedback connections between successive levels of representation; and (c) the correspondence of each representational level with a distinct processing stage (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). Thus these models suggest a high level of interaction between various stages of representation.

Other models, which will not be discussed further here, include statements about the content of retrieval cues, such as the compound-cue model (Ratcliff & McKoon, 1988), as well as more general accounts of cognitive architecture, such as distributed network models (McClelland & Rumelhart, 1985). In sum, several strong theoretical frameworks exist, which when considered jointly can account for most of the empirical data on semantic priming.

2.2.2 Neuroscientific underpinnings of semantic processing

Language comprehension occurs very fast and seemingly without effort. Given the speed at which the relevant processes operate, it is of considerable interest to examine language understanding 'online'. This involves studying neural responses with a sufficiently high temporal resolution, enabling an investigation of the processes directly related to the speech input. Several neuroscientific methods, notably Electroencephalograms (EEG, see Methods for further details) and Magne-

[4]E.g. target color when presented visually, voice gender when presented auditorily, or affective valence, when target words are of emotional content (for an overview of paradigms see McNamara, 2005).

toencephalograms (MEG) are ideally suited to adress such questions. At present we will focus more on EEG and in particular ERP studies, which have been put to very good use in providing the basis for explanatory accounts of language comprehension, reconciling conflicting evidence generated by behavioural data (Friederici, 2002). The present review of the neuroscientific data will focus explicitly on semantic aspects of speech comprehension. This will entail the functional interpretation of a salient electrophysiological marker of semantic processing, the N400 (Kutas & Hilyard, 1980; Kutas & Hillyard, 1984), as well as consider some of the attempts to localise these processes in the brain, by means of neuroimaging studies.

2.2.2.1 The N400

The N400 is an electrophysiological response first reported in a study measuring ERPs to semantically anomalous words (Kutas & Hillyard, 1980), such as the word socks in the sentence “He drank his coffee with milk and socks”. The name is derived from its deflection, which is negative, and peak latency, which is around 400 ms. Typically, the N400 is distributed broadly over the scalp, but most strongly over centro-parietal electrodes (Kutas & Federmeier, 2000). Its amplitude has been shown to correlate inversely with the word’s rated ‘cloze-probability’ (i.e. proportion of subjects rating the word as the most likely semantic completion of a phrase or sentence; Kutas & Hillyard, 1984). Since then it has been used as an electrophysiological marker of the kinds of processes described in the section on semantic priming above.

Even though the N400 is especially large to semantically anomalous sentence material, it appears to be the brain’s normal response to the continuous semantic integration occurring in speech comprehension (van Petten & Kutas, 1991). In this study it was shown that an N400 was elicited by each semantically expected word in a sentence, the amplitude of which decreased with the increasing number of words. Therefore the N400 has been argued to reflect the semantic predictability within a local as well as a global context.

The N400 has been argued to reflect constraints imposed by several levels of context, such as the immediate information contained in the context of the sentence, as well as word associations (van Petten, 1993) and global discourse-level constraints (van Berkum, Hagoort, & Brown, 1999). As a result, the N400 has frequently been interpreted as reflecting contextual integration, which emphasizes the relationship between an eliciting item and its context.

Apart from immediate contextual effects, it has also been argued that the N400 reflects the ease with which information is accessed from long-term memory (Kutas & Federmeier, 2000). Evidence in favour of this idea comes from studies investigating the sensitivity of the N400 to long-term semantic representations such as category membership (Kounios & Holcomb, 1992; Kutas & Federmeier, 2000). Thus, the N400 appears to predominantly reflect two types of memory processes: a working memory process in which the incoming semantic information is fitted with the present context, as well as a long-term memory process.

The question which aspect of word recognition (i.e. access, selection, retrieval) the N400 actually reflects is still under debate. Even though it may even be possible that the N400 reflects all three stages of word recognition, studies on the influence of depth of processing on the N400 (Chwilla, Brown, & Hagoort, 1995) have concluded that this ERP most likely reflects lexical integration rather than access, which has been echoed in subsequent studies and reviews (van Petten, Coulson, Rubin, Plante, & Parks, 1999; Friederici, 2002). Most importantly for the present research purposes, the N400 is a reliable indicator of semantic processing and can be used as such to inform on what is capable of communicating semantic information. So far, there is no evidence to date that musical targets can elicit an N400 (Kutas & Federmeier, 2000). As will become apparent in the next chapter, this fact has been efficiently exploited in a study on music processing, providing part of the conceptual edifice on which the present research is based.

2.2.2.2 Localizing semantic processing

Several studies have attempted to localise the generators of the N400 by means of both intracranial recordings as well as source localizations using MEG and EEG. These have suggested a number of generators (Nobre & McCarthy, 1995; Halgren et al., 1994; Halgren, Baudena, Heit, Clarke, & Marinkovic, 1994; Meyer et al., 2005), such as anterior medial temporal lobe structures (Nobre, Allison, & McCarthy, 1994; Nobre & McCarthy, 1995), comprising the hippocampus, the superior temporal suclus and additional frontal areas (Halgren et al., 1994, 1994), left planum temporale, superior and inferior temporal sulci, temporal pole and rhinal sulcus, insula and prefrontal cortex, as well as the right orbital cortex (Halgren, Raij, Marinkovic, Jousmäki, & Hari, 2000). A most recent MEG study reported generators of the N400 in the left inferior frontal gyrus (BA 47 and BA 45), as well as the anterior parts of the superior (BA 22) and inferior (BA 20/21) temporal gyri bilaterally (Maess, Herrmann, Hahne, Nakamura, & Friederici, 2006). This impressive variety of generators seems to depend largely on the type of stimulus material, the methods and the mathematical models used. Other methods with a higher spatial resolution have been able to shed some light on the distribution of cortical networks underlying semantic processing. As will become apparent below, particularly findings on activations in frontal and temporal regions of the brain map well onto what has been reported from attempts to localise the N400.

Studies on semantic processing using brain imaging techniques, such as functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET) typically involve observing which areas are involved in processing semantic relationships established either by word pairs or a sentence context. Usually, activations are calculated by contrasting semantic implausibility against semantic plausibility, to see which areas are involved in processing semantic aspects of language.

As indicated above, several studies provide evidence for the involvement of temporal structures in semantic processing. A study by Kuperberg et al. (2000) reported activations in the right superior and middle temporal gyry, with semantic violations compared to correct sentences. Another study, using word pairs, have reported increased activations of the left anterior medial temporal cortex (Rossell, Price, & Nobre, 2003). Several other studies report temporal lobe activity during semantic anomaly processing, either distributed bilaterally (Friederici, Rüschemeyer, Hahne, & Fiebach, 2003; Ni et al., 2000) or lateralized to the right (Kiehl, Laurens, & Liddle, 2002; Newman, Pancheva, Ozawa, Neville, & Ullman, 2001; Rodd, Davis, & Johnsrude, 2005; Rissman, Eliassen, & Blumstein, 2003). In an attempt to tie together the divergence of different processes (e.g. phonological, syntactic, semantic) during speech comprehension, recent reviews (Hickok & Poeppel, 2000, 2007) have attributed particular importance to temporal structures, part of the so-called *ventral stream* of speech processing. Particularly posterior middle-temporal and inferior temporal regions are argued to be involved in the linking of incoming phonological and existing semantic information, the so-called *lexical interface*. It is argued however that most presumably semantic information is widely distributed throughout the brain, and that these areas are more involved in mapping the sound to its meaning, rather than hosting a lexicon of all possible meanings.

In addition to the temporal areas that have been found to be engaged in semantic processing, frontal regions have also been shown to be preferentially involved, in particular the left inferior frontal gyrus (IFG; Dapretto & Bookheimer, 1999; Bookheimer, 2002; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Friederici et al., 2003; Kuperberg et al., 2000; Newman et al., 2001; Schirmer et al., 2004). Some specific hypotheses have been proposed with regards to the role of the left IFG, such as the selection of a semantic candidate among competing alternatives (Thompson-Schill et al., 1997). Others have suggested that the left IFG plays a prominent role in semantic retrieval rather than selection (Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). Recent findings however suggest a certain parsimony within the evidence, proposing posterior portions of the left IFG to be involved in lexical selection, and anterior portions to mediate the retrieval of lexical-semantic representations (Gold et al., 2006).

In sum, posterior medial and superior temporal areas, also known as the ventral processing stream, as well as the left inferior frontal gyrus appear to be involved in processing semantic aspects of the language input. Differences in the activations between studies on semantic processing are most presumably due to differences, such as the stimulus material, the modality of presentation and the task given to the participants.

2.3 Summary

The stages involved in successful word recognition, which is the basic unit of language comprehension, involve access, selection and retrieval of words stored in lexical memory. Each word activates associated words, priming their subsequent processing. Electrophysiological findings have identified a unique component reflecting semantic processing, the N400. Attempts to localise this component with source modelling have indicated a variety of different neural generators, but in conjunction with neuroimaging methods, both inferior frontal and particularly temporal brain regions appear to play a crucial role in subserving semantic processing.

Chapter 3

Music processing

As argued in the first chapter, music has frequently been seen as a non-verbal means of communication (Meyer, 1956; Bernstein, 1976; Sloboda, 1986; Swain, 1997). The present chapter will begin by giving an overview of how this communication has been operationalised empirically and focus on the properties intrinsic to music, rather than what the composer or the performer would like to communicate. A recently proposed model of music processing will be introduced with emphasis on how it accounts for musical communication. It will be suggested that it can make sense to talk of meaning in music and after a brief definition of what qualifies and what does not qualify as meaning in music, several pathways to musical meaning will be considered. Some empirical evidence proposed in favour of the notion of musical meaning will be reviewed, from which more specific predictions will be made with regards to the present thesis.

3.1 A model of music processing

The cognitive mechanisms underlying music listening and production have attracted increased scientific interest over the last twenty years. Listening and making music is an astoundingly complex feat for humans, engaging virtually all cognitive processes, such as basic perception, attention, emotion, memory, social cognitive processes as well as action (for a review, see Koelsch & Siebel, 2005). Thus, apart from an intrinsic interest in the phenomenon of music, the study of music cognition stems from its representing an ideal tool to study various cognitive processes, in isolation and interaction. Recently, this interest has spilled over to the neurosciences, which have endeavoured to investigate the neural correlates underlying the perception, cognition and action of music and music-making with considerable success (Zatorre & Peretz, 2001; Avanzini, Faienza, Minciacchi, Lopez, & Majno, 2003; Avanzini, Lopez, Koelsch, & Majno, 2005; Peretz & Zatorre, 2005). The advantages of neuroscience over traditional psychological methods apply to music processing as well. Like language, music is a temporally unfolding phenomenon and therefore the online-processes are of particular interest. On the basis of the data obtained from EEG and neuroimaging data a model has been proposed (Koelsch & Siebel, 2005, see Figure 3.1). This model provides a thereotical backbone as well as empirical predicitions with regards to meaning in music and will now be introduced.

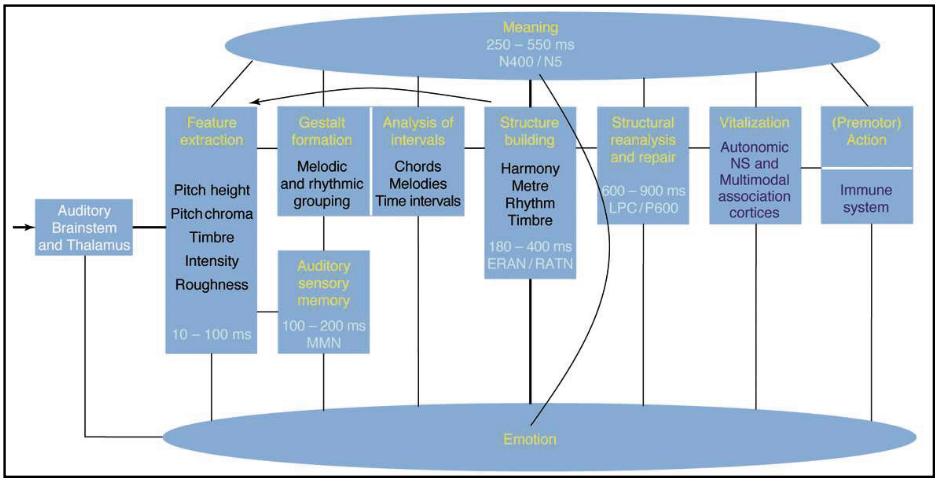

Figure 3.1: A model of the time-course of music processing by Koelsch and Siebel (2005)

Based on ERP-evidence, the model (see Figure 3.1) divides music processing into early and late stages after intial acoustic processing in the auditory brainstem and thalamus. Early processing stages involve the extraction of basic auditory features from the acoustic signal (e.g. pitch height and -chroma, timbre, intensity and roughness), and the formation of *auditory Gestalten* (e.g rhythmic and melodic grouping), which are then stored in auditory sensory memory and subjected to interval-analysis (e.g. rhythmic, melodic, harmonic). Later processing stages include the building of musical structure using the available acoustic information as well as the repair of any structural violation that may have occurred in the music. Subsequent processing stages refer to what has been termed *vitalization* and the planning of related actions. Most importantly, it is argued that each of these processing stages is somehow linked to the communication of meaningful signals, as well as emotion-related processes, be it the recognition of an emotion expressed by the music or the feeling of an emotion in response to the music. Additionally, emotional processes are argued to link to meaning in music and vice versa.

Even though based on data, much of the model by Koelsch and Siebel (2005) is still hypothetical, making various predictions, particularly with regards to what is capable of communicating meaning in music. Importantly, the model stresses an explicit link between meaning and emotion in music, as well as between meaning and all levels of music processing, which is an aspect explored further below.

3.2 What does *musical meaning* mean?

Whether music has meaning and what exactly this meaning entails has been an intensely debated topic amongst musicologists and neuroscientists alike (Lerdahl & Jackendoff, 1983; Jones & Holleran, 1991; Raffman, 1993; Sloboda, 1986; Swain, 1997; Koelsch et al., 2004; Janata, 2004). The present section will specify what is meant by meaning, how this applies to music and in what sense it is possible to talk of meaning in music.

A general misconception seems to be that words have inherent meaning, which was neither taught as such nor learned as such. Even though certain words may be more suited to describe certain objects or events better than others (e.g. onomatopaeia), these instances are exceptions and not the rule. Whether, if music were to be used as primary means of expression from birth, it would be just as effective as language, is a hypothetical question one may therefore wish to ponder. It may be that the set of discrete units of one system (e.g. phonemens, syllables, words, phrases) may be better suited to communicate efficiently than the other (e.g. pitch, melody, harmony, timbre). This line of argument however suggests that meaning is not reserved exclusively for language, but for any other system, which is learned and taught in the way that language is. Music normally has no intention to communicate anything specific, but that does not imply that it would not be possible to do so.

However most scholars argee on the basic claim of Lerdahl and Jackendoff (1983) that whatever music may mean, it is in no sense comparable to linguistic meaning, since there are no musical phenomena comparable to sense and reference in language. This is reminiscent of what Bertrand Russel once remarked of a dog's bark, which may be very eloquent, but cannot tell you that his parents were poor but honest. However at the same time, listeners, in spite of their limited knowledge of what the word *meaning* may denote in a philosophical sense, would hesitate to say that music is incapable of meaning anything, a sentiment shared by scholars (Sloboda, 1986). Thus, music may still mean something, but lack the specificity inherent to meaning in language.

A brief but relevant mention has to be made at this point that in this dissertation the term *semantic* will not be used in relation to the kind of meaning that can be conveyed by music. It is a term that has been thoroughly appropriated by the field of linguistics and certain criteria have to be fulfilled before anything can qualify as semantic. There are some who claim that nothing outside of language can have semantic meaning. Therefore, without neither wanting to prematurely terminate a discussion of meaning in music due to unsuitable terminology nor venture forth blindly into an epistemological minefield, meaning in music is predominantly discussed only by using the bare term *meaning*.

Generally, theories on meaning in music have made the distinction between referential and non-referential meaning (Meyer, 1956; Swain, 1997). Since the work presented in this thesis builds directly on both of these, they will be discussed in turn, considering some of the preceding theoretical work as well as the empirical data for each.

3.2.1 Referential meaning in music

Referential meaning appears to be what most seem to have in mind, when talking of musical meaning, for it matches the type of meaning that is instantiated by language, whereby a particular sound will be associated with a specific meaning. Music however lacks discreteness. By now, most scholars agree that the theory on emotions expressed by music espoused by Cooke (1959) and outlined in the first chapter, does not bear up in the light of the vast discrepancy of listeners' responses to music. Recently it has been proposed that the meaning music is capable of expressing should be seen in terms of a *semantic range* (Swain, 1997). Indeed language can be seen in a similar way, where for instance the word *deck* can have over six different meanings. Therefore, one has to accept that faithfulness to propositional meaning ought not be a prerequisite for the meaning expressed by music. If anything, specific musical pieces can mean a range of things. This has to be borne in mind, when discussing the scope and limits of musical meaning.

The kind of referential meaning that music is capable of communicating has been suggested to arise through several routes (Swain, 1997; Koelsch et al., 2004): (i) musical features mimicking features of real-life objects or events (e.g. high pitched undulating and playful piano sequences to resemble fresh water, as in Ravel's *jeu d'eau*); (ii) extramusical association (e.g. national anthems or advertising jingles); and most obviously (iii) the suggestion of a particular mood or emotion. The evidence for these three pathways to meaning in music has rested predominantly on neuroscientific and some behavioual evidence (Koelsch et al., 2004; Sollberger, Reber, & Eckstein, 2003).

In the study by Koelsch et al. (2004), musical pieces were chosen such that they were strongly reminiscent of a specific concept, by virtue of association or feature mimickry. Concepts included words with either concrete (e.g. needle, river) or abstract (e.g. hero, solitude) meaning. These pieces were then presented, subsequent to which either the expected target word occurred or an unrelated alternative word. Taking EEG-recordings and analysing ERPs it was found that unrelated words elicited a considerably larger N400 than the related words, indicating that music is cabable of priming semantically meaningful concepts. A language condition which was run in the same experimental session indicated that the N400 in response to the music stimuli was indistinguishable from the N400 elicited in response to sentences. This was taken as evidence that music can communicate meaning as effectively as language can.

Given that the emotional content of the musical pieces was controlled for, mood suggestion could not be counted among the routes to meaning tested in the study, in spite of its being intuitively the most obvious one. Another recent behavioural study however reported some relevant evidence and suggested promise for further research (Sollberger et al., 2003). By means of an affective priming study, it was shown that single chords which were either pleasant or unpleasant sounding, could influence the subsequent processing of emotionally valenced words, indicating that very basic musical features were already capable of signifying emotions and associated semantic processing. Thus particular musical features could effectively refer to perceptual or semantically related features of the associated concept. In sum, there is already some evidence in place suggesting that certain musical features are capable of consistently communicating the same meaning.

3.2.1.1 Some remarks on the link between meaning and emotion

Intuitively, what music seems best able to express is a wide range of emotions. In the current theoretical and empirical treatment of the subject, this will be discussed jointly with the suggestion of a musical mood, even though these might be conceptually distinct. As argued above, emotional expression has been hypothesized to represent a unique pathway to establishing meaning for the listener. So far, and as pointed out by Peretz and Coltheart (2003), the perception of an emotional expression and the actual emotional experience in response to the music have not been differentiated empirically. At this point we can only speculate with regards to their relationship. This conceptual fuzziness however does not alter the possibility that the expression of emotion in music may be meaningful. The link between emotion and meaning in music is particularly relevant to the experiments conducted in the present thesis and several lower level musical features will be explored with regards to their ability to communicate emotions and therefore meaning.

3.2.2 Nonreferential meaning in music

This appears to be the type of musical meaning that most musicologists feel safest with, for it does not make any specific claims as to whether musical meaning can function referentially like in language. The idea that music contains no semantic value but the tautological one of referring to itself dates back to the famous 19th century musicologist Eduard Hanslick (1854/1986) and has found its most vociferous champion in Stravinsky (1936), who claimed that "music is, by its very nature, essentially powerless to *express* anything at all, whether a feeling, an attitude of mind, a psychological mood, a phenomenon of nature, etc... . Expression has never been an inherent property of music" (p. 48). While unwilling to claim that music is entirely devoid of meaning, this meaning is not referential, but entirely abstract and composed of the signs established during the course of a musical piece. The continuous variations of musical themes throughout a piece of music would be a suitable example, whereby a theme may be increasingly imbued with meaning with each variation (e.g. the lover's in Berlioz's *Symphonie fantastique*).

Meyer (1956) wrote on the meaning derived from musical structure and in particular the role of tension and resolution patterns, which occur in virtually all music. Meyer argues that single musical entities do not necessarily symbolise anything (even though they can), but rather that they point to a musical consequence. The kind of meaning of a musical event is therefore borne out of its implicit suggestion of a number of possible subsequent musical events. These possibilities are constrained by the expectations, which have been implicitly learned and are subject to particular rules and hierarchies impinging on the perceptual system (Tillmann, Bharucha, & Bigand, 2000). Tightly bound up with affective responses that listeners have in response to the manipulation of their musical expectations, this theory holds that such non-referential aspects of music are meaningful to the listener familiar with the musical system.

In spite of these explicit formulations, there is no evidence to date supporting these ideas. By means of exploiting principles of general cognitive architecture (i.e. the sharing of neural resources, also outlined in the following chapter) the present thesis also involves a study directly testing the idea of non-referntial meaning in music.

3.2.3 ERPs of music processing

One way of finding out whether music is capable of conveying meaning is to investigate the neural correlates of music processing. As already discussed in the previous chapter, the N400 has been taken as the classic marker for semantic processing (Kutas & Hillyard, 1980). Any study which can show that music can elicit an N400 would therefore indicate that music can convey meaning[5]. Attempts to respond to this question date back over 20 years. Besson and Macar (1987) for instance used several conditions (e.g. sentences, geometrical forms, scale notes of increasing or decreasing frequency, and well-known tunes) and presented them in an expected as well as an unexpected condition. Only the semantically unexpected sentences

[5]Even though the N400 has also been shown to reflect other types of processing as well (Frisch & Schlesewsky, 2001).

elicited an N400, whereas the non-linguistic deviances elicited late positivities. On the basis of these findings it was concluded that under the basic and most ecologically valid conditions (i.e. well-known songs being played out of tune), no N400 can be elicited and therefore these do not constitute semantic violations.

A study by Besson and Faita (1995) played familiar and unfamiliar melodies to both musicians and nonmusicians in a normal version as well as with diatonic and non-diatonic violations. The main finding that was reported was a late positive component, taken to reflect the processing of these violations. However another component which was displayed but not explicitly reported was a negativity between 200-400 ms for violations of familiar melodies. A most recent study replicated this effect (Miranda & Ullman, in press) showing an early and posteriorly distributed negativity in response to violations of long-term musical representations. These effects can be argued to reflect the processing of violations of long-term musical representations. In terms of function, the section on the N400 in the previous chapter already shows that the N400 is sensitive to both short-term contextual integration as well as long-term representations, a framework which can account for the present data. However, this is still no evidence that the N400 in this case reflects the processing of violations of meaning.

Apart from melodic processing, the processing of harmonic expectancies has also been investigated (Janata, 1995; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Koelsch, Gunter, Friederici, & Schröger, 2000; Regnault, Bigand, & Besson, 2001), where chords were used to set up a harmonic context, which was then either harmonically confirmed or violated, according to the rules of Western harmony. Janata (1995) found two positive components, peaking at 310 ms (P3a) and another at 450 ms (P3b), elicited by the harmonic expectancy violation. The P3a increased incrementally with the degree of the violation. Seeing that the paradigm asked for overt responses, these components presumably reflect an overlap of perceptual, cognitive and decisional processes. No N400 was elicited however, by these violations of harmonic expectancy.

In a study comparing language and music processing, Patel et al. (1998) found that harmonic expectancy violations elicited a larger P600. As a result of comparable findings in the language condition (using a syntactic violation) it was concluded that the harmonic context is understood as musical structure, the violation of which is structural in nature and not semantic (hence reporting a P600 and not an N400).

The study also reports a right anterio-temporal negativity (RATN), which was argued to also reflect the processing of music-syntactic rules.

Koelsch et al. (2000) also used a harmonic expectancy violation paradigm using Neapolitan chords as violations of the context. These chords, even though in themselves consonant and perfectly acceptable, represent a very unexpected harmonic turn within a key sequence, especially when preceded by a dominant chord, which usually leads back to the tonic. The study reported two components in response to Neapolitan chords, compared to tonic chords, which have been replicated by numerous follow-up studies (for a review, see Koelsch & Siebel, 2005), an early right anterior negativity (ERAN) peaking around 180 ms and a late bifrontal negativity peaking around 500 ms (N500). Whereas the ERAN was taken to reflect the processing of a violation of harmonic expectation, the N500 was argued to indicate the increased effort required to integrate the violation into the exisiting harmonic context. As a result of its time-course and scalp distribution, the ERAN has been interpreted as reflecing music syntactic processing (Koelsch et al., 2000). Evidence with regards to the syntactic function of the ERAN is by now largely uncontested (Koelsch, Gunter, Wittfoth, & Sammler, 2005), however the functional specificity of the N500 remains an open empirical issue.

A study by Regnault et al. (2001) reported a similar effect to Janata (1995), whereby violations of the expected harmonic function elicited a P300. However, participants were also required to signal the detection at the event of interest, meaning that other potentials may have been masked by response-related ERPs. A more recent replication of the design (Poulin-Charronnat, Bigand, & Koelsch, 2006), removed this confound and reported late bifrontal negativities highly reminiscient of the N500.

Thus, there are a variety of ERP components that have been found for the processing of several types of musical violations. Whereas none of the rule violations reported an N400, but a variety of other components, only the violations of long-term musical representations could elicit a component similar to the N400. However, in this case the N400 does not seem to reflect the processing of violations of musical meaning at least not in the sense defined above, but rather the violation of long-term memory representations of musical pieces. This brief review has shown that most musical contexts do not convey meaning, as indicated by an absence of the N400. Whereas this does not suggest that music does not have meaning, it strongly implies that what kind of ERP is elicited and what musical function is being adressed depends on the nature of the context and the nature of the probe stimulus.

3.2.4 Summary

The model by Koelsch and Siebel (2005) predicts that all stages of music processing are capable of expressing emotions, eliciting emotional responses and being meaningful to the listener. In spite of some general evidence that music can convey meaning, there is no support for specifically hypothesised routes, such as via emotions or non-referential meaning. ERPs in response to music processing suggest that traditional musical contexts do not elicit an N400, but rather other components indicating all kinds of processes (e.g. perceptual, decisional, structural). As well as providing some crucial definitions of what kind of meaning one may expect to encounter when listening to music, this chapter outlined several gaps in the literature, which in spite of elaborate theoretical accounts have not been filled and which this thesis aims to adress.

Chapter 4

Shared neural resources between music and language

There is a tradition of considerable standing among musicologists and those interested in the study of music perception, to make analogies to linguistics and language perception (Bernstein, 1976; Lerdahl & Jackendoff, 1983; Aiello, 1994; Besson & Friederici, 1998; Besson & Schön, 2001; Besson & Friederici, 2005). This has ranged from a basic descriptive level, including the components that language and music consist of, to a functional level, such as how the components can be arranged to achieve some kind of a purpose (syntax, semantics, emotion). Such a comparative approach gave rise to the question whether processing music and language may be subserved by the same neural resources. This has now been debated for some years and evidence both in favour and against has been gathered. Recent theories on the comparability of music and language have tried to bring some consensus into the disparate findings (Patel, 2003). This chapter will provide an overview of the theory proposed to account for the nature of shared neural resources between music and language and present some of the evidence in favour of the idea. This provides the theoretical backbone to an empirical investigation on shared resources dedicated to processing meaning.

4.1 The controversy

The neuroscientific findings on whether the underlying neural correlates of language and music processing are comparable has been somewhat contradictory. On the one hand lesion studies list a series of well-documented cases of dissociations between language and music processing (Peretz & Kolinsky, 1993; Peretz et al., 1994; Ayotte, Peretz, & Hyde, 2002; Luria, Tsvetkova, & Futer, 1965; Marin & Perry, 1999), suggesting no significant overlap between the two domains[6]. On the other hand, EEG-studies have demonstrated considerable similarity between components elicited by harmonic expectancy violations (Patel et al., 1998) and those ERPs elicited by syntactic violations (Osterhout, Holcomb, & Swinney, 1994), such as the P600. Additional neuroimaging evidence, obtained from MEG and fMRI suggests that musical rule violations are processed in areas typically associated with language rule violation processing, such as Broca's area and its right hemispheric homologue (Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch et al., 2002). As it stands, there is a lack of consensus as to whether music and language share processing characteristics (e.g. brain structures or functional ERPs). A recent model on shared syntactic processing, proposed by Patel (2003) has attempted to resolve some of these contradictions and will be discussed below.

4.2 The shared syntactic integration resource hypothesis (SSIRH)

Patel (2003, p. 674) likens language and music to each other in that they are both "...human universals involving perceptually discrete elements organized into hierarchically structured sequences". As a result of these basic similarities it is argued that music and language can serve as suitable foils for each other to study the underlying mechanisms of complex sound processing. The central topic of Patel's review is the role of syntax in language and music, which according to him may be defined as a set of principles governing the combination of discrete structural elements into sequences. Syntactic rules are the combinatorial principles with which the incoming perceptual input is organized and of which the listener possesses im-

[6]This interpretation of the data has gone as far as proposing a separate processing module dedicated entirely for music (Peretz & Coltheart, 2003).

plicit knowledge (Sloboda, 1986; Tillmann et al., 2000; Chomsky, 1965; Saffran, Aslin, & Newport, 1996; Saffran, Senghas, & Trueswell, 2001). Whereas applying syntactic rules in language enables the transformation of linear input into a hierarchical sequence, which provides an understanding of "who did what to whom", in music, these rules entail the tracking of tension-resolution patterns subject to the rules of Western harmony (Lerdahl & Jackendoff, 1983).

To provide a parsimonious account of the contradictory data presented above, Patel (2003) proposed his shared syntactic integration resource hypothesis (SSIRH). This hypothesis draws a distinction between syntactic representations (e.g. grammatical categories and functions) and syntactic processes, which refer to the operations (e.g. integration) conducted on prior knowledge for the purpose of building coherent percepts. It is argued that representations and processes, henceforth referred to as *dual-system* are subserved by different cortical regions (Caplan & Waters, 1999; Ullman, 2001). The assumption is that language and music processing are served by the same types of mental operations (e.g. integration), but the units which are processed are not represented in the same way (e.g. word class vs chord functions). Specifically, it is argued that syntactic integration is a process required for both music and language processing.

This theory has received increasing scientific attention by providing several empirically testable predictions as well as laying the foundation for further conceptual work on shared operations between language and music.

4.3 Empirical evidence

Several studies have investigated the hypothesis of shared neural resources between music and language processing. Some recent behavioural evidence has been able to provide support for the SSIRH, by means of observing the relative influence of music material on the processing of simultaneously presented language material. In these studies, the syntactic manipulation in the language material occurred by means of varying the distance between dependent words[7], rather than explicit syntactic violations (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2007). Participants sung sentences varying in syntactic complexity to melodies that did or did not contain an out-of-key note on the last word of the relative clause. An additional

[7] This amounts to higher syntactic complexity.

condition controlled for the attentional resources dedicated to the deviant event. Comprehension accuracy of the sentences was taken as evidence for processing difficulty. As expectd comprehension accuracy was lower for sentences with higher syntactic complexitiy than lower syntactic complexity. Most importantly however, this effect was larger when melodies contained an out-of-key note, which was not the case for the control condition. These results strongly suggest that some aspect of structural integration in language and music relies on shared processing resources.

A study by Slevc, Rosenberg, and Patel (2007) used garden path sentences as syntactic anomalies in the language condition, while hearing chord sequences containing more and less expected progressions. The dependent variable was self-paced reading of the sentences and it was found that syntactically unexpected words were read more slowly compared to expected words, and that this difference was even more marked when an unexpected chord was heard at the same time, compared to an expected chord. A control condition using semantic anomalies did not show such an effect, indicating that only the syntactic resources appear to be recruited by both the music and language material.

Evidence for the SSIRH has also been obtained from an ERP study (Koelsch et al., 2005), which looked at the influence of harmonic expectancy violations on ERPs in response to various types of language stimuli (Koelsch et al., 2005). Harmonic expectancy violations have been interpreted as reflecting musical syntax violations (Koelsch et al., 2000) and it was hypothesised that if syntactic processes operate in a domain-general fashion (i.e. similarly for language and for music), then processing syntactic violations in language ought to be influenced by the simultaneous processing of harmonic expectancy violations. To test this, subjects were presented with both language and music material simultaneously. The visually presented language material was either syntactically incorrect (gender violation), semantically unexpected (with low-cloze probability), or correct, whereas the harmonic chord sequences ended either on an expected Tonic or an unexpected Neapolitan chord. Participants were told to focus on the acceptability of the language material. Analysis of the ERPs for each of the language conditions revealed an N400 for the semantically unexpected sentences and a left anterior negatitivty (LAN) for the syntatcic violation (a component reported for syntactic gender violations; Gunter, Friederici, & Schriefers, 2000). Most importantly, the LAN was significantly reduced when a harmonically unexpected chord was presented at the end of the harmonic sequence, whereas the N400 was unaffected by the degree of expectedness in the music material. A control experiment showed that this LAN reduction was not due to mere deviance detection mechanisms, as a simple auditory mismatch did not affect the component. It was therefore concluded that syntactic operations in language and music share neural resources, as indicated by the reduction in the LAN when presented simultaneously with the harmonic expectancy violation. This is therefore an important piece of evidence providing direct support for the assumption that language and music share neural resources for certain types of operations.

4.4 Summary

In sum, there appears to be substantial evidence in support of the notion that language and music do share some processing resources when it comes to syntactic operations. Several further questions come to mind, such as if other cognitive processes shared by music and language also recruit the same neural resources, as is indeed surmised by (Patel, 2007). The present overview serves as the basis for extending the framework to semantic processing, and provides part of the theoretical foundation for Experiment 4.

Chapter 5

The role of expertise and musical training

When learning how to play an instrument, one engages in a wide variety of different tasks, such as motor learning, visuo-, auditory- and motor association and coordination, auditory memory and spatial orientation, as well as honing all the skills engaged in increased exposure to the Western musical system. There is by now ample evidence of experience-driven neuroplasticity manifesting itself on a behavioural and neural level. The necessary question arising out of this, is whether the increased musical activity leads to significant changes with regards to the processing of certain aspects of the musical input. Answers to this would indicate the extent to which certain skills can be trained or not and whether a basic competence of musical skills is present in the majority of the population. The present chapter will provide a brief description of the term *plasticity* and its supposed underlying mechanisms. After an outline of empirical accounts of the neuroanatomical and functional changes resulting from increased musical training (usually instrumental), differences between musically trained and untrained subjects at the behavioural level (including cognitive and emotional aspects) will be discussed.

5.1 Plasticity

It has been known for some time that cortical representations in adult animals are not fixed and static entities, but much rather continuously changed by experience.

The cortex is capable of preferentially allocating an area to represent whatever peripheral input source proportionally most used. Changes in these cortical representations seem to underlie learning tasks, which depend on the use of the peripheral input that they represent. It has been emphasised that the mechanisms underlying cortical plasticity, may be the more elementary aspects of synaptic plasticity, such as long-term potentiation (LTP) and long-term depression of excitatory postsynaptic potentials (EPSP; Buonomano & Merzenich, 1998). Thus synaptic plasticity and cortical plasticity have been linked, in that the former has been argued to underlie cortical map reorganisation.

The neural mechanisms underlying cortical representational remodelling are thus surmised to result from synaptic plasticity, most importantly from LTP of excitatory synapes following Hebbian learning rules. Hebbian plasticity simply refers to the increaed synaptic strength resulting from two neurons firing simultaneously (Hebb, 1949). First examples of Hebbian plasticity were described in CA1 neurons in the rat hippocampus, whereby the presynaptic input from CA3 axons was paired with postsynaptic depolarization leading to subsequent enhancement of EPSP amplitude (Kelso, Ganong, & Brown, 1986). This enhancement was seen to be long-lasting and described as LTP and confirmed one of Hebb's primary postulates that the extent of pre- and postsynaptic activity modulates the extent of synaptic connection. A particular subtype of the glutamate receptor, namely the NMDA receptor, has been argued to fulfill the role of coincidence detector between pre- and post-synaptic firing. These receptors allow for an influx of calcium CA^{2+}, which is a critical step in the induction of LTP (Lamprecht & LeDoux, 2004). Calcium in turn serves as intracellular messenger, activating additional signalling pathways, which activate enzymes and which in turn initiate gene transcription and protein synthesis. This chain of events has been argued to lead to structural alterations at the postsynaptic membrane, stabilising excitatory synaptic transmission over longer periods of time (Buonomano & Merzenich, 1998).

This brief review on the underlying mechanisms of cortical plasticity will now provide a basis for considering the reasons and evidence for plasticity in musicians. Below a brief description will be given of the potential for cortical plasticity in musicians followed by some of the evidence obtained in recent years. This will then be followed by a review of how these changes may be reflected in both behavioural and cognitive performance, by virtue of skill-specific performance tests.

5.2 The special case of musicians

Musicians offer an ideal model in which to study the plastic changes associated with experience (Münte, Altenmüller, & Jäncke, 2002). Musical training is an intensely demanding feat, requiring several hours of instruction each week and several hours of practice each day and for those training at conservatoire level this lasts over a considerable time period. Performing music at a (near-to) professional level is possibly one of the most complex human accomplishments including demanding bimanual coordination and high-level motor production, often at breakneck speed. Apart from the fine motor skills which are learned and enhanced while learning an instrument (including the synchronization with the musical output of others during joint music-making), musicians often receive explicit instructions in music theory and history, as well as refining perceptual auditory analysis, so called *aural training*. It is therefore to be expected that, given this extensive and long-term training of a wide range of motoric, cognitive, behavioural and social domains, one ought to see dependent changes in the brain, both neuroanatomically and functionally. In addition, given these associated changes in cortical organization one would also expect this to be reflected in improved performance in both musical and possibly even non-musical tasks and skills, the latter of which is known as so-called *transfer effects*. The present section will discuss the findings accumulated with regards to both brain- and associated behavioural changes resulting from musical training.

5.2.1 Neuroscientific evidence

The neuroscientific evidence in support of plasticity resulting from musical training has largely been obtained from studies looking at neuroanatomical differences between musically trained and untrained subjects, which typically entails measuring the extent of differences in gray or white matter. Alternatively, the functional differences have also contributed significantly to our understanding on the recruitment of differing cortical structures depending on musical training, whereby differences in brain responses, as measured by EEG/MEG and fMRI/PET, to mostly musical stimuli inform on the degree of cortical reorganization. It must be added that unless these group differences are correlated with some behavioural performance, mere cortical differences are not very informative. After all, the change in cortical organization ought to be reflected either qualitatively or quantiatively in behavioural measures.

5.2.1.1 Neuroanatomical perspectives

Even though post-mortem studies of brain anatomy have a long tradition, recently high-resolution fMRI has enabled the study of anatomical details *in vivo*. An area, for which one would expect anatomical differences as a result of increased training, is the motor cortex. In a study by Amunts et al. (1997), the size of the primary hand motor area was estimated by determining the intrasulcal length of the posterior bank of the precentral gyrus. Musicians showed greater intrasulcal length on both sides, but more so on the right, non-dominant hemisphere. This was reflected in reduced asymmetry scores in musicians, which in turn correlated strongly with the time at which musical training began, suggesting cortical reorganisation, rather than a predisposition to become a musician.

An additional structure focused on to investigate the neuroanatomical differences between musicians and nonmusicians also included the corpus callosum (CC). The CC is of particular interest, as it represents the connectedness between the two hemispheres and presumably increased bimanual coordination, which is required of instrumental musicians and which ought to be reflected in increased fibre crossings between the hemispheres. It was found that musicians who commenced their musical training before the age of seven had an increased anterior CC compared to non-musicians, which consists predominantly of the crossing of motor fibres (Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Lee, Chen, & Schlaug, 2003).

The cerebellum is a structure required for the precise timing of movements. This particular skill is not just required of any musician performing a piece of music, but also of joint music-making, where the precise synchronisation of motoric output is essential. A recent study demonstrated that paticularly male musicians had greater mean relative cerebellar volume than male non-musicians, which could not be ascribed to difference in the overall brain volume (Schlaug, 2001).

One additional study could demonstrate that musicians showed increased gray matter in perirolandic regions, primary and secondary somatosensory areas, premotor areas, anterior superior parietal areas, as well as left inferior temporal gyrus, cerebellum, Heschl's gyrus, and IFG (Gaser & Schlaug, 2003b, 2003a). These are all areas that have been somehow implicated in music perception or production (Gaser & Schlaug, 2003b, 2003a) and in concurrence with findings from other laboratories (Schneider et al., 2002, 2005) this demonstrates that musical training modifies brain areas, reflecting the operation of cortical reorganisation. In sum, there is considerable evidence for anatomical differences in a variety of brain areas involved in motor and auditory processing.

5.2.1.2 Functional perspectives

Functional studies have investigated changes in cortical response to various types of stimuli, usually auditory, as a result of musical training. These are either taken between two groups of subjects, one with and the other without musical training, or within the same group of subjects receiving musical training over a longer period of time, comparing responses before and after the training.

For instances, in a seminal study, somatosensory evoked magnetic fields in string players were investigated (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995). Larger cortical representations of the digits of the left hand (with which string players do the fingering) were revealed in musicians compared to nonmusicians. This was not the case for the right hand, suggesting specific reorganisation as a result of the increased finger training. In addition, this was found to correlate with the onset of musical training. The first studies to reveal significant changes in receptive function showed that the evoked magnetic fields in response to piano tones compared to pure tones, was around 25 % stronger in musicians compared to nonmusicians (Pantev et al., 1998). In a follow-up study investigating trumpet and violin players, cortical responses were strongest for each musician's own type of instrument (Pantev, Roberts, Schulz, Engelien, & Ross, 2001).

Several ERP studies have shown that cortical responses to sequenced auditory stimulation also differ between musicians and nonmusicians. For instance, the mismatch negativity (MMN) is known as a marker of pre-attentive change detection in regular auditory sequences (for a review, see Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). Recent studies have shown that the MMN signalling detection of very slight sound omissions, is significantly greater by musicians compared to nonmusicians (Rüsseler, Altenmüller, Nager, Kohlmetz, & Münte, 2001). Additionally, musicians showed an MMN to slightly impure dissonant chords in a string of pure consonant chord sequence, which was absent in nonmusicians (Koelsch, Schröger, & Tervaniemi, 1999). Other components known to reflect the processing of musical structure, such as the ERAN (Koelsch et al., 2000), have also been shown to differ between musicians and nonmusicians, whereby the former group displayed a larger ERAN to harmonic irregularities than the latter (Koelsch, Schmidt, & Kansok, 2002). This difference in processing harmonic expectancy violations was replicated in an fMRI study (Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005), where musically trained adults showed stronger activations of the inferior frontal gyrus (BA44) bilaterally for irregular chords at the end of five-chord sequences, compared to nonmusicians. Other studies have shown the involvement of the IFG in the processing of harmonic expectancy violations (Koelsch et al., 2002; Maess et al., 2001; Tillmann, Janata, & Bharucha, 2003; Tillmann et al., 2006) and this suggests that musical training enhances the involvement of structures subserving harmonic processing.

Further fMRI studies have suggested several cortical networks to be differentially involved between musicians and nonmusicians during various tasks. Hund-Georgiadis and von Cramon (1999) for instance compared performance and activation between musicians and nonmusicians during a complex finger-tapping task. Whereas for both groups the performance improved during the experiment, the activation of different brain regions was shown to accompany this improvment. Learning in musicians was shown to recruit the contralateral primary motor hand area, secondary motor areas, as well as premotor, somatosensory and cerebellar areas. In nonmusicians, learning was accompanied by increased activations of ipsi- and bilateral primary and secondary motor areas. A further study showed increased activations in musicians when merely listening to previously played melodies, as well as when playing these melodies on a mute keyboard, in areas known for auditory-somatosensory integration (e.g. dorsolateral and inferior frontal gyrus, superior temporal gyrus, supramarginal gyrus, as well as supplementary, and primary motor areas). These activations were not present for nonmusicians, showing the presence of a shared network between auditory and motor processing in musicians only.

In sum, there appears to be considerable evidence from both neuroanatomical as well as functional studies on the cortical reorganization associated with learning to play an instrument. The fact that in most of these studies, these changes are associated with the onset at which instrumental training began, strongly suggests that these changes are a result of, rather than a prerequisite to musical training.

5.2.2 Behavioural evidence

In a recent review paper on the role of musical experience in music perception and cognition, Bigand and Poulin-Charronnat (2006) urge for an increased understanding of which kind of musical capacities are the result of implicit learning processes and which ones are the result of musical experience. The present evidence on implicit learning of musical structure (Tillmann et al., 2000) is contrasted with the findings of brain changes associated with music training, posing the question what behavioural changes actually occur as a result of musical training. The present section will give a summary of the findings on cognitive tasks measuring performance on musical structure processing, as well as reports of emotional responses to music by both musicians and non-musicians.

5.2.2.1 Cognitive processing

The difference in cognitive performance as a result of musical training has been assessed using a variety of musical aspects. One of these has focussed on the processing of and sensitivity to the underlying structure of musical pieces. This was tested by comparing the ability to assign a particular varition to its theme (Bigand, 1990), after having manipulated either surface features (e.g. rhythm, melodic contour, pitch range) or deep features (e.g. harmonic progression). It was found that even though musicians performed significantly better than nonmusicians (72% compared to 58%), the latter group still performed above chance. This suggests that even though nonmusicians are worse at processing deep structural features than musicians, they are still sensitive to more abstract musical information and use it to process musical pieces.

An additional and possibly the most explored aspect of music cognition is the processing of musical expectancies. As already argued in chapter 3, musical regularities, such as the rules of Western harmony, are learnt by mere exposure to music in everyday listening situations (Tillmann et al., 2000). For instance the rules underlying harmonic structure processing can be explained by something known as the *hierarchy of stability* (Bharucha & Krumhansl, 1983), whereby in a given tonal contex, certain tones (Krumhansl & Shepard, 1979) or chords (Krumhansl & Kessler, 1982) will sound more stable or final than others. This perceived stability correlates directly with the harmonic distance between the home key and the event in question, as displayed by the Circle of Fifths. Thus chords closer to the home key will sound more stable and less stable the further these chords move away. Early studies show that the operation of harmonic rule learning leads to certain expectancies with regards to harmonic progressions (Bharucha & Stoeckig, 1986; Bharucha & Stoeckig, 1987). Perceptual judgements on target chords are made faster when the chord is harmonically closely related to the preceding prime chord than when distantly related. This has been referred to as the *harmonic priming effect*, which has also been the subject of several studies, particularly with regards to processing advantages of musically trained over nontrained subjects (for a review see, Bigand & Poulin-Charronnat, 2006). As a result of the considerable musical input, it is very plausible to predict that musicians should be more skilled at or influenced by harmonic distances. However, several studies indicate that even though musicians are generally faster at responding to target chords in general, the size of the harmonic priming effect did not differ significantly between the groups (for a detailed review see, Bigand & Poulin-Charronnat, 2006; Bigand, Poulin, Tillmann, Madurell, & D'Adamo, 2003; Bigand, Tillmann, Poulin-Charronnat, & Manderlier, 2005).

Another aspect of music processing that has been tested is the learning of new musical idioms. Given the increased sensitivity to certain acoustic properties, it may be easier for musicians to extrapolate the unknown rules of musical input. A recent experiment looking at rule learning of serial music, directly compared musicians and nonmusicians (Bigand, Perruchet, & Boyer, 1998). After a brief learning phase in which fourty canons specifically composed to embody the transformational rules, which serial music is subject to were presented, rule learning was tested by asking participants which of two canons similar in surface structure was composed in a similar fashion to the ones presented in the learning phase. It was found that even though the groups performed relatively well, given the complexity of the task (musicians at 62.10% and nonmusicians at 60.71% accuracy), this difference did not reach significance. Thus abstract rule learning appears to be one of basic musical skills, which does not benefit from musical training.

5.2.2.2 Emotional processing

An additional aspect that has been investigated with regards to potential differences between musicians and nonmusicians is that of emotional perception of and responses to music. As such, no clear hypothesis can be made whether musical training should lead to changes in emotional responses. Emotional responses occur widely in the human population and are not restricted to a privileged few. Nonetheless, this issue was adressed in a recent study on characterizing the emotional content of musical pieces (Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005). It was found that when classifying the same musical pieces, both groups came up with the same number of emotionally distinct groups, into which the pieces could be classified and overall the responses beween the groups correlated very highly. A replication of the same experimental procedure was carried out, but decreasing the length of musical pieces to 1 second. The responses between the two groups still did not differ, showing that emotional classification does not benefit from musical expertise.

Another very recent study demonstrated that musical training had no impact in the emotional response to harmonic expectancy violations, as measured by psychophysiological indicators (Steinbeis, Koelsch, & Sloboda, 2006). In sum, there appears to be no evidence that musicians may respond emotionally differently to music than nonmusicians.

5.2.3 Summary

The neuroscientific data suggest that there are significant differences in the neural representation of auditory and musical stimuli as well motoric functions resulting from musical training. Very often these changes appear to correlate with the onset at which the participants started to learn an instrument, which strongly suggests that training modified specific brain structures. The behavioural data on the other hand suggest hardly any differences between musically trained and untrained participants and is therefore at odds with the neuroscientific evidence. Are the anatomo-functional brain changes resulting from musical training meaningful, if this is not reflected in some qualitative or quantitative perceptual or behavioural benefit? One way of resolving this issue is that behavioural performance was rarely measured when the neuroscientific data was acquired. The behavioural tests listed above may merely tap into different processes, not measured by the neuroscientific studies. More importantly however, the behavioural data gives no indication of the underlying mechanisms by means of which each group arrives at the measured performance. Even though the scores may look the same, musicians and nonmusicians may have gotten there by very different means. This issue of same outcome but different mechanisms can only be resolved by simultaneous acquistion of cortical activity. Thus, quantitatively there may be no strongly discernible differences between musically trained and untrained participants in music processing, however, the qualitative differences between the groups may be considerable, and indeed scrutiny of the neuroscientific evidence strongly suggests that this is the case.

Chapter 6

Methods

6.1 Electroencephalogram (EEG)

The cerebral cortex, crucially involved in all aspects of sensory perception and higher cognition, including affect and motivation, is a folded sheet of cells varying in cortical thickness between 2 to 5 mm. The various types of cells making up the cortex have been classified into two types: *pyramidal* and *nonpyramidal*, differing with regards to morphology, laminar distribution and neurotransmitter content.

The effective recording of electrical brain activity at the scalp depends on the structural organisation of the cortex. Pyramidal cells are organised in parallel to one another and their dendritic organization occurs perpendicular to the scalp. When active, pyramidal cells produce ESPS by means of an ionic current flowing inward through the synaptic membrane and outward through the extrasynaptic membrane. The perpendicular organisation of pyramidal cells coupled with the fact that EPSPs summate well over time, produces electrical activity, which can be measured at the scalp. For this activity to be sufficiently detectable, many pyramidal cells (over 10000) are required to be synchronously active, the result of which summates at the scalp, known as an *open field* and is large enough to be recoded by an EEG (Kandel, Schwartz, & Jessel, 1991).

However, not all electrical activity from the brain can be measured at the surface of the scalp. Specifically, if neurons are not organised in a fashion which is parallel and perpendicular to the scalp, activity cannot be recorded by means of a surface EEG. This is particularly the case with subcortical structures.

At the very least an EEG requires the placement of two electrodes. An *active electrode*, placed over a site of neuronal activity and an *indifferent electrode*, placed over a site of little or no such activity. The electrical activity over the active electrode is recorded with reference to the indifferent electrode. The recorded activity can be analysed in both a temporal as well as a frequency domain, whereby activity is divided into frequency-bands denoting the rate of oscilation. The frequencies of potentials typically vary from 1- 30 Hz and the amplitudes typically range between 20 - 100 μV, which is obviously somewhat attenuated when recorded at the scalp.

One enormous advantage of EEG over other methods, notably functional imaging, is its high temporal resolution and its ability to track electrical activity continuously. As already mentioned, a disadvantage is the fact the scalp EEG cannot record neural activity which does not occur in parallel and perpendicular to the scalp or synchronously precluding activity from so called *closed field* structures.

6.1.1 Event-related potentials (ERPs)

The term event-related potentials refers to all electrocortical potentials occuring prior, during or after a motoric or psychological event. Typically derived from deducting the activity of electrodes over active sites from electrodes over indifferent sites, ERPs reflect voltage amplitude and are reported in μV. Because they are usually recorded at the scalp, amplitudes are attenuated compared to measurements obtained by means of intracranial or subdural recordings.

The recording of ERP signals occurs time-locked to a particular event. Simultaneously, there is still ongoing and spontaneous EEG activity, which is not related to the specific event under investigation and which produces so-called "noise", contaminating the overally EEG signal and making it impossible to infer changes in the ERP to the time-locked event. A necessary consequence of this poor signal-to-noise ratio derived from a single event is to increase the ratio by means of repetitive stimulation with the same event type. The ERPs are then averaged for each event type separately and as a result, any systematic electrocortical response should become evident from the improved signal (see Figure 6.1).

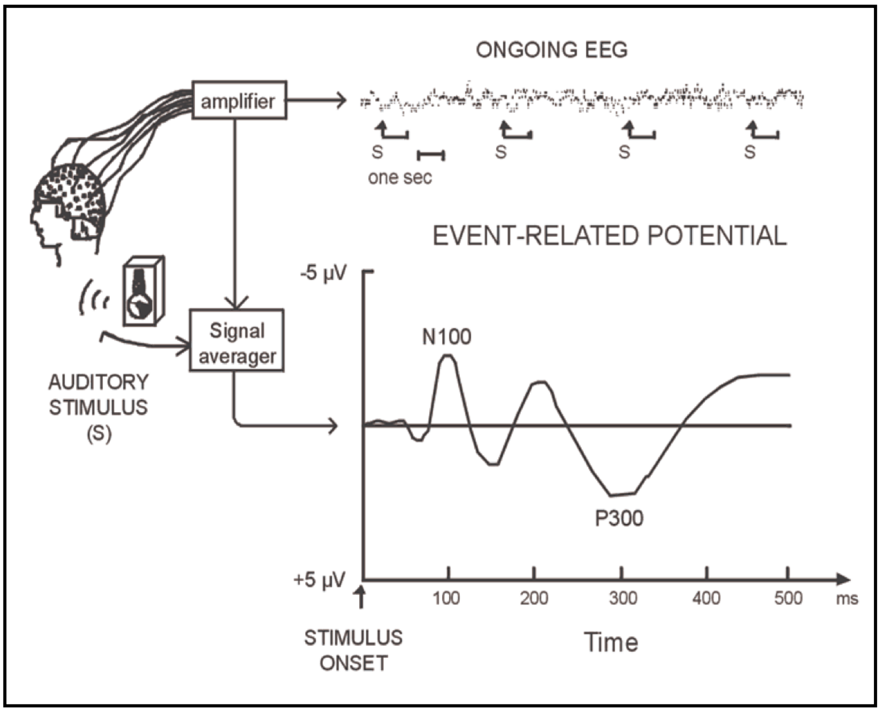

Figure 6.1: EEG recording, amplifying and averaging

ERPs occur systematically to specific events, which for empirical psychologists represent the experimental conditions, and which are described and interpreted with regards to their psychological function. Descriptions of ERP components are predominantly based on polarity, latency and scalp distribution. Onset, peak and offset of a potential can be specified with millisecond precision. The time-course is a crucial interpretational aspect to be considered with reference to the automaticity of components and their presumed cognitive function.

The function that an ERP component reflects has been divided into two main classifications: exogenous and endogenous components (Rugg & Coles, 1995). Exogenous components have been argued to reflect basic sensory or perceptual processes related to the transmission of sensory information from the peripheral sensory system to the cortex. These generally occur prior to or around 100 ms (e.g.

N100) and are largely unaffected by the current psychological state of the subject (attention, task relevance etc). By definition, endogenous components reflect the subjects' interaction with the environment (Rugg & Coles, 1995) and usually occur after 100 ms (e.g. N400). Like most distinctions the strength of a clean split prior- and post 100 ms has not borne up well. Instead the existence of an exogenous-endogenous continuum coextensive with time has been proposed (Rugg & Coles, 1995). The study of ERPs in cognitive neuroscience entails ascribing functional significance to particular components (e.g. the N400 for semantic processing).

6.2 functional Magnetic Resonance Imaging (fMRI)

This part will briefly outline some basics concerning the physical principles of Nuclear Magnetic Resonance (NMR) and proceed to describe the basic characteristics of the Blood Oxygenated Level Dependent (BOLD) response used for fMRI.

6.2.1 Principles of NMR

The most frequently occuring element in the human body is hydrogen (H), which consists between 60-70% of water. The nucleus of H consists of a single proton and rotates about its axis (precession), which is also known as spin (see Figure 6.2). This rotation reflects the inherent magnetic dipole moment of the nucleus. Outside of any magnetic field, the spin of a proton is random (see Figure 6.3). However as soon as these are placed into an external magentic field, the spin will align either parallel (low energy) or anti-parallel (high energy) to the magnetic field. The energy level is also dependent on the strength of the magnetic field. Nonetheless the proportion of anti-parallel to paralell alignment is 1 to 1.000.000.

The frequency at which the protonic spin precesses around the axis of the magnetic field (z) is also known as the Larmor frequency. The value of this frequency can be expressed as follows: $v_0 = {}_{\gamma}B_0$, where B_0 constitutes the strength of the main magnetic field and $_{\gamma}$ the gyromagnetic ratio.

NMR exploits the difference in energy between magnetic dipoles aligned parallel and anti-parallel to the magnetic field (see Figure 6.4). To bring all dipole spins into the same alignment, external energy needs to be applied to tip the dipoles. This is performed by means of electromagentic radiation, also known as the radio-

Figure 6.2: Spin of a nucleus

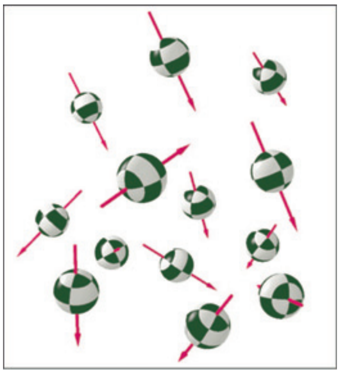

Figure 6.3: Magnetic dipoles of H protons are randomly aligned

frequency (RF) pulse. This pulse leads to an altered spin precession and subsequent orientation of the net magnetic field of H protons, M_0, which is orthogonal to the main magentic field B_0.

The new alignment of precession of M_0 is known as transverse magnetisation M_{xy}, which is reversed as soon as the RF pulse is switched off (see Figure 6.5). The tip in M_0 as induced by the RF pulse causes a signal, which can be measured by a current in a coil, and is also known as free induction decay (FID). Once the RF pulse ceases, the longitudinal magnetisation (M_0) recovers, whereas the transverse magnetisation (M_{xy}) disintegrates. These processes occur over time, known as T_1 and T_2 relaxation respectively. Each play a pivotal role in MRI measurements.

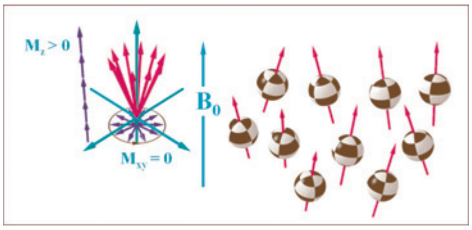

Figure 6.4: Magnetisation (M_0) is aligned with B_0

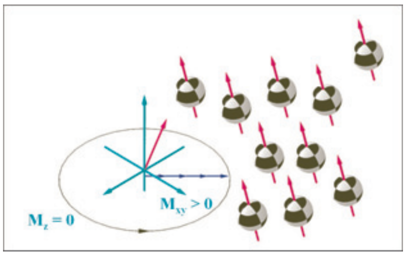

Figure 6.5: Transverse magnetisation

T_1 relaxation is characterized by protons returning from a high energy state to a low energy state and occurs as a resul of interaction between excited nuclei of interest and unexcited surrounding nuclei. The additional energy which is released over the time period T_1 is ultimately transformed into kinetic energy, such as heat and can be absorbed by the surrounding molecules. The length of T_1 is directly related to the type of tissue under investigation (e.g. fat, grey brain matter, white brain matter). Therefore the measurement of T_1 is particularly suited to acquire images of brain structures.

T_2 relaxation is characterized by a change in the phasic fluctuation of individual molecules in Larmor frequency rather than the energy state. T_2 relaxation is dominated by interactions between spinning nuclei already excited. What is most important however with regards to the two different times of relaxation is that the measurement of each gives indications of changes, such as structure or blood supply, the latter of which the meaning will be discussed below.

6.2.2 The BOLD response

Several imaging methods, particularly PET and fMRI measure metabolic or vascular paramters. The physiological basis of these methods is the fact that brain cell activity is associated with local changes in metabolism. More specifically, the level of glucose and oxygen consumption is measured by means of resulting changes in the MRI signal. The underlying assumption is that any change in cognitive demand or function ought to be accompanied by an increased need for energy, which, via so-called neurovascular coupling, is supplied by oxygen in the blood.

The reason why changes in blood supply can be measured by MRI is because blood consists of oxygenated- and deoxygenated hemoglobin, which possess varying degrees of magnetic susceptibility. Wheras oxyhemoglobin is diamagnetic and with little influence on the local magnetic field, deoxyhemoglobin is paramagnetic, leading to inhomogeneities in the local magnetic field. Generally, deoxyhemoglobin leads to an overall decrease of T_2 relaxation. Since neuronal activity requires additional oxygen, oxyhemoglobin is increased, in relation to which deoxyhemoglobin is decreased, leading to increases in T_2 relaxation and an increase in the local MRI signal. These are the underlying mechanisms of the BOLD response, which has been used by fMRI. Recent explanations of the BOLD response have argued that local field potentials may be the underlying neuronal activity giving rise to the measured signal (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). Even if such explanations have been proposed, it remains to be said that the BOLD response is merely an indirect marker of neuronal activity and any interpretations made with regards to linking activations to cognitive functions need to be done with considerable care.

6.2.3 fMRI-data preprocessing and statistical evaluation

The result of an fMRI acquisition is a time course of three-dimensional images of varying BOLD signal intensity. These images have to be processed in series of several steps, before undergoing statistical analysis. Below will give a brief outline of the rationale behind and the paramaters of each processing step taken.

Motion correction

One basic problem in the acquisition of fMRI images is its susceptibility to movement artefacts caused by head movements in the scanner. These artefacts can be corrected by applying a motion correction involving the estimation of movement relative to a reference scan (usually the first scan).

Slice timing

Seeing that the functional slices are acquired in sequential order, this results in a different acquisition time for each slice. In order to avoid any problems this might cause in the subsequent statistical analysis, all voxels of each slice are shifted backwards in time to simulate their simultaneous acquisition by means of interpolation.

Filtering

fMRI data can be contaminated by various artefacts, such as heartbeat. Some of these artefacts occur at a specified frequency, which when applying the correct filter can be removed from the signal. By applying a Fast Fourier Transformation (FFT) both noise occuring at high frequencies (low pass filter) as well as noise occuring at low frequencies (high pass filter) can be removed from the signal.

Spatial smoothing

There are several motivations for applying spatial smoothing to the acquired data. According to matched filter theorem, the optimum smoothing kernel corresponds to the anticipated size of the desired effect. The spatial scale of the hemodynamic response is, according to high-resolution optical imaging experiments, around 2-5mm, which is also suggested for fMRI experiments. Additionally, smoothing will ensure a more normal distribution of errors and ensure the validity for subsequent statistical test and inference. An additional reason is that smoothing ensures that the data represents a Gaussian field, which is required for the subsequent statistical inferences about local changes in the fMRI signal, which are made using Gaussian random field theory. Also, when averaging over several subjects, smoothing allows to project the data onto a spatial scale where homologies in functional anatomy are expressed among subjects.

Coregistration

This step is taken to ensure that the functional images acquired during the experimental session are mapped onto a precise anatomical reference. In the present evaluation package (LIPSIA Lohmann et al., 2001), this is carried out in two steps: (1) A transformation matrix is computed, registering the presently acquired 2D anatomical slices with the previously acquired 3D data set. This matrix describes a rigid, affine linear transformation using 3 translational and 3 rotational parameters, designed to find the best fit between the 2D and the 3D data sets; (2) The transformation matrix is applied to the raw functional data.

Normalisation

In order to average the acquired data over a group of subjects and to make subsequent statistical inferences on the data and particularly over specific anatomical regions, the data needs to be normalized to a standard size or anatomical space (usually Talairach & Tournoux, 1988). There are two types of normalising the data, linearly and non-linearly. During the linear transformation, the data is linearly scaled to a standard size by a rigid body transformation (with three translational and three rotational parameters). During the non-linear transformation an anatomical 3D data set (i.e. the model) is deformed such that it matches another 3D anatomical data set (i.e. the source) that serves as a fixed reference image using a deformation field.

Statistical Parametric Mapping

Statistical inference in LIPSIA (Lohmann et al., 2001) is based on a general linear model (GLM; Friston, Holmes, Poline, Price, & Frith, 1996). The GLM can be desribed as an equation that relates the observed changes in fMRI signal to the expected signal change, by expressing the observed variable as a linear combination of explanatory components and a residual error term.

Part II

Empirical part

Chapter 7

Research questions

The preceding overview has given some clear indications of where further research on meaning in music may be required. In spite of evidence that music can activate meaningful concepts (Koelsch et al., 2004), there has been no systematic research into which aspects of the musical signal are capable of doing so. Thus, the main focus of the present work is on exploring the postulation of several routes to meaning (see Chapter 3). Specifically, two of these routes will be focussed on: suggestion of musical mood/emotion and non-referential tension-resolution patterns. An additional focus of the present work addresses the comparability of neural representations of meaning in music and language. This last issue also arose out of a long-standing question in the literature on the neurocognition of music, namely whether musical targets can elicit an N400, thereby indicating semantic processing. Thus, the main research objectives of the present thesis, can be divided into three sections.

The first objective was to adress the question of whether emotion communicated by a musical stimulus is capable of influencing subsequent emotional processing on a semantic level. Because previous research has so far only provided evidence for meaning in music to arise out of extra-musical associations and music mimicking features of objects or concepts (Koelsch et al., 2004), it was considered important to examine whether the hypothesis that emotion in music is a primary means of communicating meaning can be supported or not. To this end the ability of basic lower level musical features, such as harmonic roughness, intervals and timbre to communicate meaning was investigated (Experiments 1-3 respectively). Lower level musical features were manipulated in their perceived pleasantness to set an emotional context, with which subsequently presented emotional words were either congruous or not. It was hypothesised that, should basic emotional cues in music be able to communicate meaning, then the musical context ought to exert a significant influence on the N400 amplitude. This issue was summed under the heading "Meaning and Emotion".

A second objective was to investigate whether non-referential pieces of music, such as tension-resolution patterns are capable of communicating meaning. This notion has circulated among musicologists for quite some time, but to date there is no empirical support. Under the assumption that similar cognitive processes recruit the same neural resources, it was tested whether ERPs in response to tension-resolution patterns can be modulated by simultaneously presented language material (Experiment 4). It was hypothesised that should tension-resolution patterns be capable of communicating meaning, a specific ERP in response to these should be modulated specifically by semantic anomalies in the language material. This question was summed under the heading of "Meaning and tension-resolution patterns and Meaning".

The third and final objective of the present thesis was to compare the representation of meaning in music and language. ERP correlates to language have suggested the N400 to indicate semantic processing. The aim of Experiment 5 was therefore to see whether, given the appropriate context, musical targets can also elicit an N400. The major obstacle of how to include sufficient information into a single musical stimulus was overcome by using chords varying in perceived pleasantness as targets. It was hypothesised that if single chords are represented semantically, then congruity between a chord and a preceding emotional word context ought to exert a significant influence on the N400 amplitude. A further study on comparing the representations of meaning in music and language was done by comparing the integration of word targets and chord targets in two separate blocks, using the paradigms of Experiment 1 and Experiment 5 respectively in the same session in an fMRI experiment (Experiment 6). It was hypothesised that if processing the meaning of music and language is represented in a similar fashion, then comparable neural correlates ought to be observed for the increased processing of incongruous word and chord targets. In addition, this study served to tease apart contradictory findings from Experiment 5, where an N400 effect for musically trained participants could be explained by both a response- and a semantic-conflict account. This entire research question was summed under the heading of “Comparing the neural representations of meaning in language and music”.

Both the first and third research issue were investigated with respect to musical training. As discussed previously, there are a number of reasons to assume that musical training affects the processing of musical input. The scant evidence to date on differences in processing semantic or emotional aspects of music as a result of training, suggests that this has no impact (Steinbeis, Koelsch, & Sloboda, 2005; Steinbeis et al., 2006; Bigand et al., 2005). Given the complete lack of any prior work on group differences with regards to the present research questions, the role of musical training in processing the meaning of music was treated as an open empirical issue.

Chapter 8

Experiments 1-3

Meaning and Emotion

Intuitively, emotion would appear to be the most obvious way in which music can communicate. Of all signals music contains, emotional ones are the most prevalent, regardless of whether one feels emotions in response to music or simply recognizes their expression (Juslin, 2003). Thus, by communicating an emotion, however basic, music can refer to a variety of different affective states, which are more or less unanimously understood by listeners familiar with the musical idiom (Juslin, 2003). Recent evidence even suggests that certain emotions portrayed in music, may be universally recognized, as people totally unfamiliar with Western music correctly identified the prevalent emotion expressed by a piece of music (Fritz et al., in preparation). To date, there is no evidence that emotions portrayed in music can be understood as meaningful signals. Given that the only investigation into whether music can communicate meaning explicitly controlled for the emotional content of target words (Koelsch et al., 2004), whether emotion in music can communicate meaning is still an open empirical issue.

Recent models on music processing have hypothesised that each and every step of music processing is capable of leading to some emotion-related response (Koelsch & Siebel, 2005) and that each of these can in turn activate processes related to meaning. Since these are still working hypotheses, without empirical support, the present focus is on three such musical features to test this: harmony, intervals and timbre (Experiments 1-3 respectively). By use of a cross-modal affective priming paradigm, it was tested whether single musical features varying in affective valence, can prime the processing of subsequently presented words.

Cross-modal paradimgs have been successfully employed both for studies on semantic priming (Holcomb & Anderson, 1993) as well as on affective priming (Schirmer et al., 2002; Schirmer & Kotz, 2003; Schirmer et al., 2005).

Affective priming typically entails the presentation of an affectively valenced (i.e. pleasant or unpleasant) prime stimulus followed by an affectively valenced target stimulus. Either the stimulus valence of the target matches with that of the prime (i.e. both pleasant or unpleasant) or it does not (i.e. pleasant-unpleasant; unpleasant-pleasant). Theory states that the processing of an affective target should be influenced by the valence of the preceding prime stimulus, either by facilitating matched target processing or delaying mismatched target processing. Whereas these paradigms have been primarily employed to assess the psychology of evaluative processes they have also been used to assess the general influence of affect in stimulus processing (Musch & Klauer, 2003).

There are several issues relevant for conducting an affective priming experiment, particularly the stimulus onset asychrony (SOA) and the experimental task. Findings have so far suggested that affective priming only works with SOAs at 200 ms or less (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Klauer, Rossnagel, & Musch, 1997). With longer SOAs the priming effect disappeared, from which it was concluded that the affective activations are short-lived and the resulting priming effect is due to automatic processes, rather than strategic ones (McNamara, 2005). Because the present research questions were very similar to the ones addressed by Schirmer and Kotz (2003) an SOA of 200 ms, as was used in their study, was also presently employed.

Tasks used in affective priming paradigms typically involve either the identification of a target attribute, pronunciating the target, or most frequently evaluating the target. Because it was deemed desirable to compare the results between the use of a word target and a chord target (Experiments 1 and 5 respectively), the evaluative decision task was considered the optimal choice.

In order to see whether a specific musical feature is capable of communicating affect and thus, if emotion is also a route to meaning, primes always consisted of chords manipulated either on harmonic roughness, intervals or timbre. The manipulation of each of these features has been shown to affect emotional responses (see the Introduction section of each experiment for details), which ought to transfer onto the subsequent processing of word content. Word targets were presented visually 200 ms after the onset of the prime (see also Figure 8.1). Participants had to decide whether the word target had a pleasant or an unpleasant meaning. Each word was presented twice, either matching or not matching the valence of the preceding musical prime. Dependent variables were the speed and accuracy of target word evaluation. In addition an EEG was recorded and ERPs were analysed. The primary component of interest to these analyses was the N400, as that has been shown to reflect semantic processing (see Chapter 2). Thus, if these single musical properties can convey meaning, the N400 ought to be sensitive to the match between musical prime and word target.

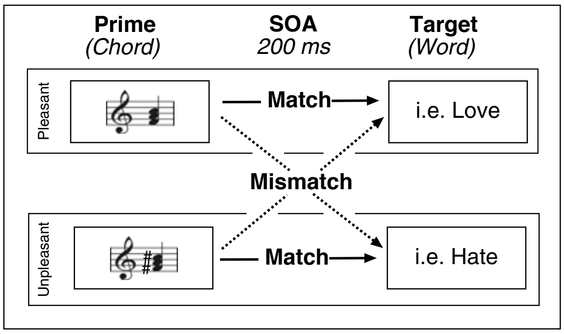

Figure 8.1: Design of the affective priming paradigm. Chords are used as primes and words as targets. In order to see whether certain musical features are capable of affective communication, chords are varied along affective dimensions of the musical feature under investigation

With each musical variation carried out in separate experiments, it was hypothesised that should the affective information contained in the acoustic parameter of a musical stimulus communicate meaning, congruent prime-target pairs should elicit a smaller N400 amplitude as compared to incongruent pairs. In addition, congruency between target and prime should also affect the response times and accuracy of target evaluation, where congruent target words should elicit faster and more correct responses than incongruent target words. The role of musical training was additionally investigated. No previous studies could provide testable hypotheses with regards to its influence on semantic processing of music, whereupon it was treated as an open empirical issue.

8.1 Experiment 1

8.1.1 Introduction

The aim of this experiment was to examine if harmonic roughness is capable of communicating meaning. Psychoacoustically, it has been suggested that the perception of harmonic roughness, specifically consonance and dissonance, is a function of the regularity of frequency ratios with which the simultaneously presented tones resonate (Plomp & Levelt, 1965). Typically, consonant music is perceived as pleasant sounding and dissonant music as unpleasant sounding, which has been shown in numerous studies. For instance, infants show a marked preference for consonant over dissonant melodies (Zentner & Kagan, 1996). Subsequent rating experiments have confirmed this and functional imaging experiments have implicated a network of brain areas also known to be involved in emotional processing. The processing of dissonance reliably activated regions known to correlate highly with perceived unpleasantness, such as the amygdala and the parahippocampal gyrus (Koelsch, Fritz, von Cramon, Müller, & Friederici, 2006; Blood, Zatorre, Bermudez, & Evans, 1999), whereas the processing of consonance elicited responses in structures dedicated to processing pleasurable events, such as the ventral striatum (Koelsch et al., 2006; Blood et al., 1999; Blood & Zatorre, 2001). This can be considered as strong evidence that harmonic roughness can modulate affective responses in music listeners.

Given the literature, one may assume that harmonic roughness contains information capable of signalling affective categories, such as pleasantness and unpleasantness, thereby communicating basic emotional information. It was therefore hypothesised that target words congruous with the harmonic roughness dependent valence of the preceding prime (consonance = pleasant, dissonance = unpleasant) should elicit a smaller N400 than incongruous target words. In addition the evaluative decision on congruous target words was hypothesised to be faster and more accurate than on incongruous target words.

8.1.2 Methods

8.1.2.1 Participants

Twenty musically trained (10 females) and 20 musically untrained (10 females) volunteers participated in the experiment. On average, musically trained participants were 23.6 years of age (sd 4.36) and musically untrained participants were 24.75 years of age (sd 2.51). Musicians had received approximately 12.3 years of musical training. All subjects were right-handed, native German speakers, with normal or corrected to normal vision and no hearing impairments.

8.1.2.2 Materials

The prime stimulus material consisted of 48 chords, of which 24 were consonant and therefore pleasant sounding and of which 24 were dissonant and therefore unpleasant sounding. The consonant chords were major chords built either on the root (in C: c-e-g-c) or on the second inversion (in C: g-c-e-g). Dissonant chords involved two types of dissonance, one using the first, the augmented fourth, and the seventh key of the scale (in C: c-f#-b-c) and another using the first, augmented second, and augmented fourth keys (in C: c-c#-f#-c). Both consonant and dissonant chords were played in each of the twelve keys of the Western musical scale, leading to 24 chords in each affective category. On average chords were 800 ms long. Chords were created in Cubase, Steinberg and modified in Cool-Edit, possessing a final sampling rate of 44.1kHz and a 16-bit resolution. A previous rating experiment (see Appendix A) conducted with both musically trained and untrained participants established that on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant, consonant and dissonant chords were significantly different from one another in their perceived pleasantness (consonant = 1.7 and dissonant = 3.9). There were no group differences in the valence ratings.

Experimental target words comprised 24 pleasant and 24 unpleasant words (see Appendix A). On average, pleasant words were 5.7 and unpleasant words 5.5 letters long. A previous rating experiment established that on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant, the affective meaning of pleasant and unpleasant words was perceived to differ significantly (pleasant = 1.7 and unpleasant = 4.4). Additionally, pleasant and unpleasant words were not found to differ in terms of the abstractness or concreteness of their content, with approximately equal number of both abstract and concrete words within and between each affective category.

For each chord, one pleasant and one unpleasant target word were chosen, which was done randomly and altered for each participant. Each chord was played twice, followed once by a congruous word and once by an incongruous word (see also Figure 8.1). There were therefore four experimental conditions: match and mismatch conditions for pleasant chords as well as match and mismatch conditions for unpleasant chords. There were 96 trials in total, with 24 pleasant match trials, 24 pleasant mismatch trials, 24 unpleasant match trials and 24 unpleasant mismatch trials. Trials were pseudorandomized and presented over two blocks of 48 trials each.

8.1.2.3 Procedure

The experiment was conducted in a sound proof and electrically shielded cabin. Participants were seated in a comfortable self-adjustable chair facing a computer screen approximately 1.2 meters away. Chords were presented from two loudspeakers positioned to the left and right of the participant. Visual targets appeared 200 ms following the onset of the chord on the screen in front. Participants were instructed to decide as fast and accurately as possible whether the meaning of the word was pleasant or unpleasant. Responses were made with a button-box, pressing left for pleasant and right for unpleasant, which was switched after the first half of the experiment. A practice run preceded the experiment and was repeated if necessary.

8.1.2.4 EEG recording and data analysis

The EEG was recorded using Ag-AgCl electrodes from 60 locations of the 10-20 system and referenced to the left mastoid. The ground electrode was applied to the sternum. In addition, a horizontal electrooculogram was recorded, placing electrodes between the outer right and outer left canthus, for subsequent removal of eye-movement related artefacts. A vertical electrooculogram was recorded placing an electrode above and below the right eye. Electrode resistance was kept below 5 kΩ and the EEG was recorded at a sampling rate of 500 Hz. The data were filtered off-line using a band-pass filter with a frequency-range of 0.25-25 Hz (3001 points, finite impulse response, fir) to eliminate slow drifts as well as muscular artefacts. To remove eye-movement-related artefacts, data were excluded if the standard deviation of the horizontal eye channels exceeded 25 μV within a gliding window of 200 ms. To eliminate movement-related artefacts and drifting electrodes, data were excluded if the standard deviation exceeded 30 μV within a gliding window of 800 ms. ERP averages were computed with a 200 ms prestimulus baseline and a 1000 ms ERP time window.

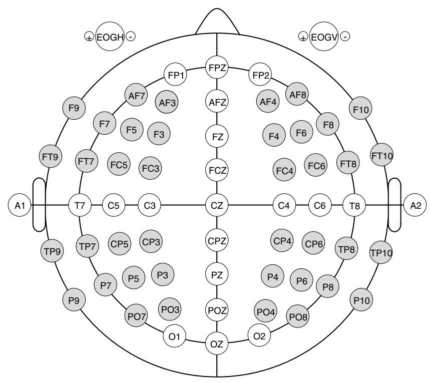

Figure 8.2: White circles indicate recorded electrodes, which were not analysed. Grey electrodes indicate the electrodes which entered the statistical analysis. Midline electrodes indicate the partitioning between the ROIs

For statistical analysis, ERPs were analyzed by repeated-measures Analysis of Variance (ANOVA) as univariate tests of hypotheses for within-subject effects. Electrodes were grouped into four separate Regions of Interest (ROI; see also Figure 8.2): left anterior (AF7, AF3, F9, F7, F5, F3, FT9, FT7, FT5, FT3), right anterior (AF8, AF4, F10, F8, F6, F4, FT10, FT8, FT6, FT4), left posterior (TP9, TP7, CP5, CP3, P9, P7, P5, P3, PO7, PO3), and right posterior (TP10, TP8, CP6, CP4, P10, P8, P6, P4, PO8, PO4). To test for specific patterns of scalp distribution, anterior and posterior ROIs established the factor AntPost and left and right ROIs established the factor Hemisphere. The time window for statistical analysis of the ERPs was 300-500 ms, based on visual inspection and those reported in previous studies (Koelsch et al., 2004).

Only correctly evaluated trials entered the statistical analysis. To test for an effect of prime valence on target processing, the factors Prime (pleasant/unpleasant) and Target (pleasant/unpleasant) were entered into the analysis separately. A significant interaction between Prime and Target was taken as an affective priming effect, indicating ERP differences between congruous targets and incongruous targets. For display purposes of ERPs, congruous and incongruous trials are depicted without differentiating further along valence and filtered with a low-pass filter of

10 Hz (301 points, fir). The additional between-subject factor Training (musically trained/musically untrained) entered all statistical analyses to test for any differences resulting from musical training.

8.1.3 Results

8.1.3.1 Behavioural results

Reaction times and accuracy were analysed separately. The factors Prime and Target were entered into a repeated measures ANOVA in addition to the between-subject factor Training. Analysis of the reaction times revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between factors Prime and Target ($F(1,38) = 26.83, p < .0001$), suggesting that both musically trained and untrained participants evaluated affectively congruous target words faster than affectively incongruous target words (see Figure 8.3). There were no further interactions or main effects.

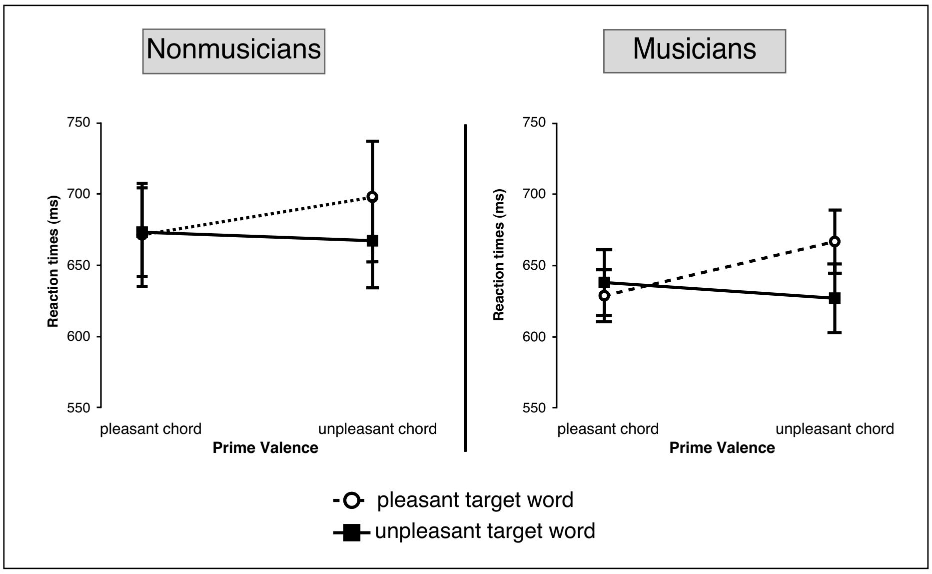

Figure 8.3: Mean reaction times (± 1 SE) for evaluative decisions on pleasant and unpleasant word targets

The analysis of performance accuracy revealed high performance of both groups (musicians: 98.4%; nonmusicians: 98.9%). There were neither significant interactions nor any main effects, showing that error rates did not differ between groups nor were they sensitive to the relationship between valence of prime and target.

8.1.3.2 ERP results

Analysis of the ERPs in the time-window of 300-500 ms revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between factors Prime and Target ($F(1,38) = 17.45, p < .001$), indicating a larger N400 for incongruous target words compared to congruous target words for both musically trained and musically untrained participants (see Figure 8.4).

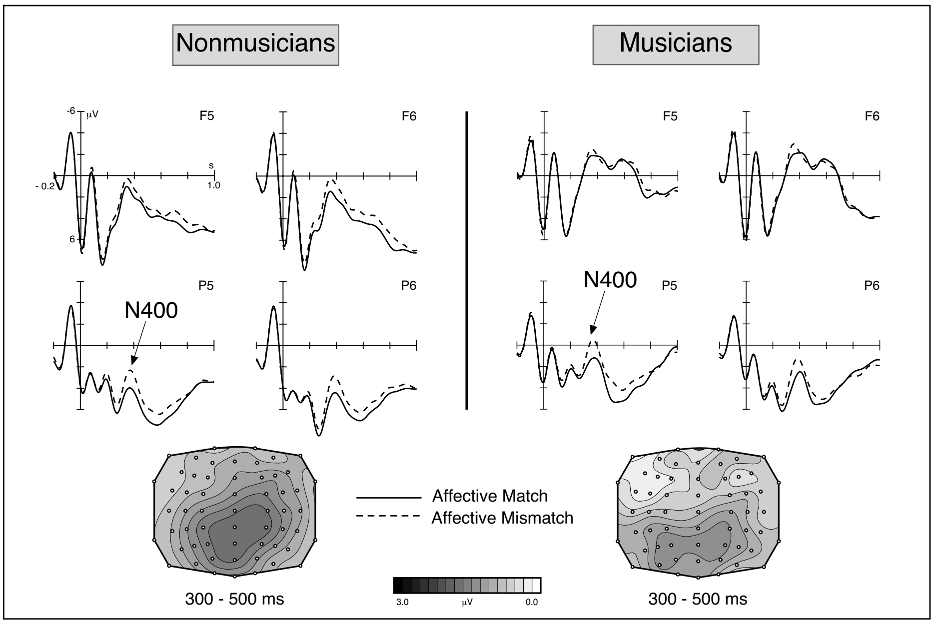

Figure 8.4: ERPs locked to onset of the target word. ERPs in response to an affective mismatch between prime and target valence (red line) resulted in a larger N400 between 300-500 ms compared to ERPs in response to a prime-target match (blue line). Both musicians and nonmusicians show the effect, which is slightly stronger over parietal electrodes.

There was also a significant interaction between the factors Prime, Target and AntPost ($F(1,38) = 5.65, p < .05$), suggesting the N400 to be larger over posterior than anterior sites. There was no further interaction with the factor Hemisphere nor any other significant main effects. For effects over more electrodes, see Figures C.1 and C.2 in Appendix C.

8.1.4 Discussion

Both musically trained and untrained participants evaluated congruous word targets significantly faster than incongruous word targets. This evident priming effect was accompanied by an increased N400 for the incongruous target words compared to the congruous target words. Seeing that the N400 has been taken to indicate semantic processing, the present findings suggest that harmonic roughness is capable of communicating affectively meaningful signals. This is therefore the first piece of evidence in support of the prediction made by Koelsch and Siebel (2005) that harmonic roughness already communicates meaningful information, which can transfer onto the processing of other meaningful concepts.

The absence of an affective priming effect for accuracy of responses can be accounted for by the very high performance of both groups. The task was therefore relatively easy and accuracy may not have been sensitive to the congruency of prime-target pairs (for similar reasoning, see Schirmer & Kotz, 2003).

The fact that this effect could be observed for both musically trained and untrained participants suggests that expertise does not modify the processing of affectively semantic properties contained in basic features of the auditory input. This appears to be in line with some previous findings, where both musicians and non-musicians were equally able to correctly classify the emotion of a musical piece, based on no more than 1 second of the music (Bigand, Filipic, & Lalitte, 2005).

Several mechanisms have been proposed to account for the various behavioural priming effects found in the literature (for reviews, see Neely, 1991; McNamara, 2005; Musch & Klauer, 2003), such as spreading activation, expectancy-based priming and semantic matching. The first of these is a mechanism argued to operate automatically, whereby the representation of each entry in the mental lexicon is connected to words closely related in meaning. The activation of one entry will automatically spread to activate closely related words. Compared to both expectancy-based priming, whereby subjects generate a list of words possibly connected to the prime, and semantic matching, whereby the subject scans preceding information when coming across a new item, and which both are therefore highly controlled processes, spreading activation is fast-acting, of short duration and does not require attention or awareness (Shiffrin & Schneider, 1977). Affective priming typically functions only at SOAs at or below 200 ms, which has also been argued to reflect the automatic nature of affective processing and evaluative decisions (Musch & Klauer, 2003). Thus, spreading activation would appear to be a likely mechanism, which can explain the observed priming effects. Hence, even though it is unlikely that single chords possess a lexical entry, they may still activate affective representations, which spread onto affectively related representations, in this case affective target words. More recently, a Stroop mechanism has also been proposed to account for affective priming effects, whereby the accrued prime and target information is integrated into a weighted sum, the weights of which are subject to both automatic and strategic processes, which lead to the response time of the decision (Wentura, 2000; Musch & Klauer, 2003). A recent empirical test of the extent to which response tendencies may account for affective priming effects (Wentura, 2000) has shown that even though judgmental tendency models can account for a considerable amount of the affective priming data, it is unlikely that the mechanisms underlying affective priming and semantic priming are fundamentally different from each other. Much rather affective priming appears to be the result of several mechanisms at work, including spreading activation (shared by semantic priming) and judgmental tendencies (shared by Stroop tasks). The ERP data recorded in this experiment yields important insights with regards to this and is discussed below.

At present also no clear answer can be given, whether the observed behavioural priming effect is the result of facilitated responses to the matched target words, an inhibitory effect to the mismatched target words, or a mixture of both. It is well documented that inhibition is small or nonexistent for SOAs shorter than 300 ms (de Groot, 1984; Neely, 1977; McNamara, 2005), which indicates that the present priming effect is the result of facilitated responses on matched target words. Even though an investigation using neutral target words would be required to adequately respond to this issue, it has been shown that the use of these in affective priming studies does not always constitute a reliable baseline (Schirmer & Kotz, 2003).

The use of an evaluative decision task implies conflict at the response level in addition to the one at the level of affective meaning, since the evaluative decision requires giving incongruous responses for incongruous trials (Wentura, 2000). Thus incongruous targets represent a mismatch on a higher level (affective meaning) as well as a lower (response) level. Whereas it cannot be ruled out that the observed behavioural effect can be accounted for by a basic response conflict/tendency explanation, the neural data suggests that this alternative account cannot be supported. ERP studies of stimulus-response conflict typically report a fronto-central N200 component (340-380 ms), presumably generated in the caudal anterior cingulate cortex (cACC; van Veen & Carter, 2002). The present data however suggest a distinctly different ERP component, both by virtue of its latency as well as its distribution, which is highly reminiscent of the N400 typically found for semantic violations. Thus, even though the behavioural effect may in part be explained by stimulus-response conflict, the underlying neural processes still suggest that harmonic roughness is a feature of the musical input, which is capable of both communicating affective meaning and priming the processing of subsequently presented affective words.

8.2 Experiment 2

8.2.1 Introduction

The aim of this experiment was to see if harmonic intervals can also communicate affective meaning as has been hypothesised (Koelsch & Siebel, 2005). In this experiment, the term harmonic intervals is used to refer to the subtle difference between a major and a minor chord. The difference between the two is the third step of the scale, which in the minor key is one semitone lower than the major key, the detection of which presumably requires a fine-grained analysis of pitches and their intervals.

The experimental literature on a link between harmonic intervals and emotion has a long history. An early study showed that major pieces of music are classified more often as happy than music pieces in a minor key and minor pieces more often as sad than major pieces (Hevner, 1935). Recently, this has found further empirical support in studies designed to investigate whether emotion conveyed by music is determined most by musical mode (major/minor) and tempo (slow/fast; Gagnon & Peretz, 2003). Using the same set of equitone melodies, participants had to judge whether they sounded happy or sad. It was found that musical mode was a highly significant predictor for listeners's judgements, with major melodies being rated significantly more often as happy than minor melodies, which in turn were rated significantly more often as sad, than major melodies. Additional evidence has shown that this ability to perceive emotions in music dependent on mode, is already present in children aged 6 to 8, but not younger ones (dalla Bella, Peretz, Rousseau, & Gosselin, 2001).

There is therefore considerable evidence in favour of harmonic intervals to communicate affective meaning such as happiness and sadness. It was therefore hypothesised that target words congruous with the musical mode dependent valence of the preceding prime (major chord = happy, minor chord = sad) should elicit a smaller N400 than incongruous target words. In addition, the evaluative decision on congruous target words was hypothesised to be faster and more accurate than on incongruous target words.

8.2.2 Methods

Participants

Twenty musically trained (10 females) and 20 musically untrained (10 females) volunteers participated in the experiment. On average, musically trained participants were 22.7 years of age (sd 3.82) and musically untrained participants were 22.88 years of age (sd 2.44). Musicians had received approximately 12 years of musical training. All subjects were right-handed, native German speakers, with normal or corrected to normal vision and no hearing impairments.

Materials

The prime stimulus material consisted of 48 chords, of which 24 were in a major key and therefore happy sounding and of which 24 were minor and therefore sad sounding. Major chords were built either on the root (in C: c-e-g-c) or on the second inversion (in C: g-c-e-g). Minor chords were also built either on the root (in C: c-e-flat-g-c) or on the second inversion (in C: g-c-e-flat-g). Both major and minor chords were played in each of the twelve keys of the Western musical scale, leading to 24 chords in each affective category. Chords were created in Cubase, Steinberg and modified in Cool-Edit, possessing a final sampling rate of 44.1kHz and a 16-bit resolution. On average chords were 800 ms long. A previous rating experiment (see Appendix A) conducted with both musically trained and untrained participants established that on a scale of 1 to 5, where 1 meant happy and 5 sad, major and minor chords were significantly different from one another in their perceived happiness (major = 2.5 and minor = 3.6). There were no group differences in the happiness ratings.

Experimental target words comprised 24 words with a happy and 24 words with a sad meaning (see Appendix A). On average, happy words were 5.4 letters and sad words 5.3 letters long. A previous rating experiment esatblished that on a scale of 1 to 5, where 1 meant happy and 5 sad, the affective meaning of happy and sad words was perceived to differ significantly (happy = 1.6 and sad = 4.4). Additionally, pleasant and unpleasant words were found to not differ in terms of the abstractness or concreteness of their content, with approximately equal number of abstract and concrete words in each affective category.

For each chord, one happy and one sad target word were chosen, which was done randomly and altered for each participant. Thus each chord was played twice, followed once by a congruous word and once by an incongruous word. There were therefore four experimental conditions: match and mismatch conditions for happy chords as well as match and mismatch conditions for sad chords. Thus there were 96 trials in total, with 24 happy match trials, 24 happy mismatch trials, 24 sad match trials and 24 sad mismatch trials. Trials were pseudorandomized and presented over two blocks of 48 trials each.

The Procedure and ERP Recording and Data analysis were the same as in Experiment 1.

8.2.3 Results

8.2.3.1 Behavioural results

Reaction times and accuracy were analysed separately. The factors Prime and Target were entered into a repeated measures ANOVA in addition to the between-subject factor Training. Analysis of the reaction times revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between factors Prime and Target ($F(1,38) = 12.11, p < .001$), suggesting that both musically trained and untrained participants evaluated affectively congruous target words faster than affectively incongruous target words (see Figure 8.5). There were no further interactions or main effects.

The analysis of performance accuracy revealed high performance of both groups (musicians: 97.7%; nonmusicians: 96.2%). There were neither significant interactions nor any main effects, showing that error rates did not differ between groups nor were they sensitive to the relationship between valence of prime and target.

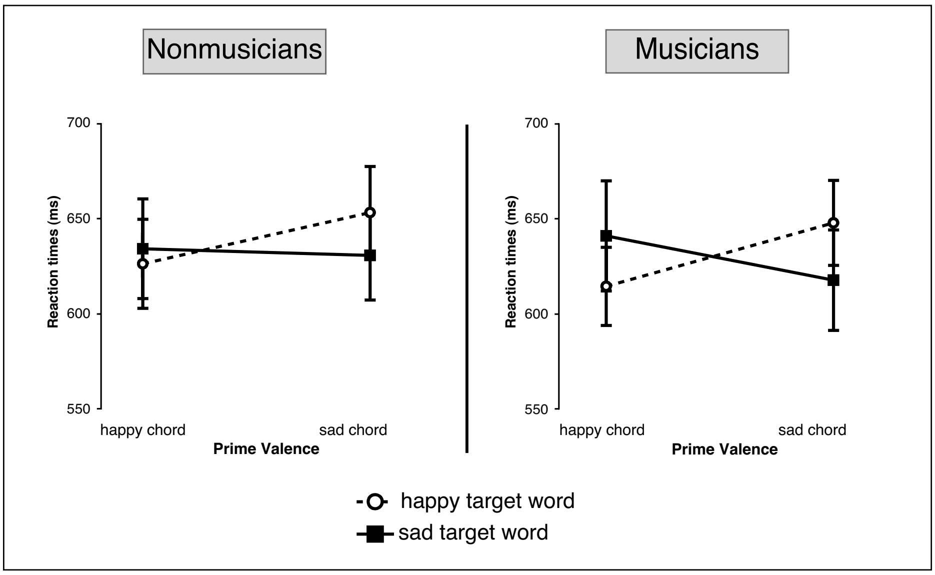

Figure 8.5: Mean reaction times (± 1 SE) for evaluative decisions on happy and sad word targets

8.2.3.2 ERP results

Analysis of the ERPs in the time-window of 300-500 ms revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between factors Prime and Target ($F(1,38) = 33.69, p < .0001$), indicating a larger N400 for incongruous target words compared to congruous target words for both musically trained and musically untrained participants with a broad scalp distribution (see Figure 8.6). There were no further interactions with any of the other factors, nor any significant main effects.

In spite of suggestive visual evidence of further ERP effects in later time windows, these could not be statistically confirmed. For effects over more electrodes, see Figures C.3 and C.4 in Appendix C.

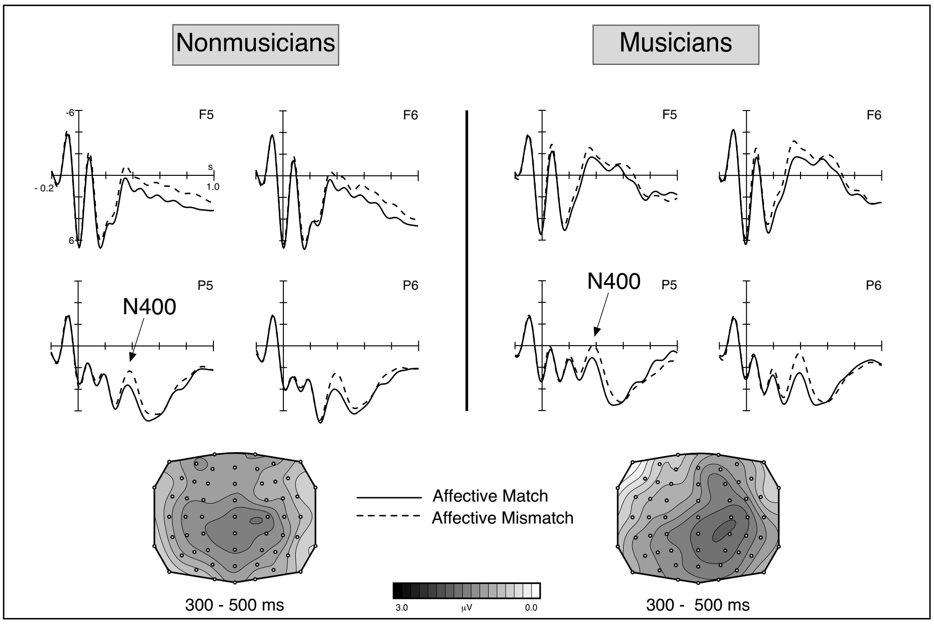

Figure 8.6: ERPs locked to onset of the target word. ERPs in response to an affective mismatch between prime and target valence (red line) resulted in a larger N400 between 300-500 ms compared to ERPs in response to a prime-target match (blue line). Both musicians and nonmusicians show the effect, which is broadly distributed over the scalp.

8.2.4 Discussion

Both groups of participants showed a significant behavioural priming effect, which was accompanied by a larger N400 for target words mismatched in valence to the preceding chord prime. These findings strongly suggest that musical mode is capable of communicating affectively meaningful signals. It is striking that as little as the manipulation of one semitone (the difference between major and minor chords) is sufficient to communicate affective meaning. This shows that the analysis of harmonic intervals is linked to establishing meaning in music and lends empirical support for the hypotheses outlined in the model by Koelsch and Siebel (2005), proposing a link between the analysis of harmonic intervals and meaning in music.

This effect was observed for both musically trained and untrained participants, which provides further evidence to that obtained in Experiment 1 suggesting that expertise does not modify the processing of affectively semantic properties contained in basic features of the auditory input.

There was no priming effect found in the accuracy of responses, which fits with the data reported in Experiment 1. Similarly, the task may have been too easy for accuracy scores to be sensitive to the prime-target relationship.

It may be argued that this experiment is merely a replication of Experiment 1, because the manipulation of a harmonic interval automatically entails a manipulation of harmonic roughness. A previous rating experiment however indicated that whereas major and minor chords were perceived as differing on the happy/sad dimension, they were not rated as significantly different on the pleasant/unpleasant dimension (see Appendix A). Thus, unless one is to discount the explicit ratings given by participants, this can be taken as evidence that even though harmonic roughness was manipulated, the manipulation of a semitone suggested a different or perhaps additional affective meaning to that conveyed by harmonic roughness (i.e. happy/sad as opposed to pleasant/unpleasant), which is supported by the differing scalp distributions between the N400 reported in this experiment and Experiment 1. This in turn suggests that the present experiment provides evidence for subtle differences in harmonic intervals to be capable of communicating affective meaning not mediated via harmonic roughness.

Like in Experiment 1, the most likely mechanism to account for the presently observed affective priming effect and the associated N400 is spreading activation. Equally, it is highly unlikely that the observed effect is a result of mere conflict at the response level resulting from the evaluative decision task used.

8.3 Experiment 3

8.3.1 Introduction

The aim of this study was to see if instrumental timbre is also capable of communicating affective meaning, as has been hypothesised (Koelsch & Siebel, 2005). So far, there is virtually no literature on a relationship between instrumental timbre and emotion. This may have partly to do with the fact that timbre has been difficult to define empirically. Definitions of musical timbre have commonly been more in terms of what it is not, rather than what it is[8]. Generally, it seems to have been agreed on that timbre is the tonal colour or texture that allows one to distinguish the same note played by two different instruments.

By use of mulitdimensional scaling (MDS) a variety of psychoacoustic parameters have been identified to correlate with the perception of different timbres. These have included attack time of the sound, as well as various spectral parameters, such as the spectral centroid, which measures the average frequency of a spectrum, weighted by amplitude, the spectral flux, a measure of the change within a frequency spectrum over the duration of the signal, as well as the spectrum fine structure, which measures the attenuation of even harmonics of the sound signal (McAdams, Winsberg, Donnadieu, Soete, & Krimphoff, 1995). A most recent study has shown that in dissimilarity ratings attack time, spectral flux and spectrum fine structure are used most saliently to differentiate between different timbres (Caclin, McAdams, Smith, & Winsberg, 2005), however more work is required in order to fully understand which parameters are used for the perceptual analysis of timbre. In sum, the perception of timbre evidently seems to depend on more than one psychoacoustic property.

There is some evidence in the literature on prosody that the timbral quality of vocal utterances (e.g. distribution of energy in frequency spectrum) is strongly related to the communication of an emotion. In a study which set out to test to what extent listeners can infer emotions from vocal cues, Banse and Scherer (1996) recorded the vocalization of 14 different emotions as portrayed by professional actors, which were then classified blindly by judges and underwent a psychoacoustic analysis. Specific psychoacoustic patterns could be identified with each vocalized emotion.

[8]Timbre is argued to refer to those aspects of sound quality *other* than pitch, loudness, perceived duration, spatial location and reverberant environment (McAdams, 1993; Menon et al., 2002).

It could be shown that the expression and perception of anger, fear, sadness and joy partly depend on the relative energy in the high- vs low-frequency spectra, which is known to express vocal timbre.

To date, the only empirical link between the perception of emotion in musical timbre is a study on the mismatch negativity (Goydke, Altenmüller, Möller, & Münte, 2004). Violin timbres differing in emotional expression (happy or sad) were used to create "emotional" standards and deviants. It was found that emotional deviants elicited an MMN, which was interpreted by the authors as the fast and accurate perception of emotional timbre. However, the study confounded perceptual differences and emotional expression, since happy and sad timbres differed in terms of basic perceptual features. Thus, apart from the ratings taken prior to the experiment, this does not constitute a clear piece of evidence that musical timbres can communicate emotions.

Even though there is no direct evidence on the peception of emotions expressed in musical timbre, the work on timbre in vocal productions suggests that this feature of the acoustic input may be more generally capable of expressing emotions. It was therefore hypothesised that target words congruous with the timbre dependent valence of the preceding prime should elicit a smaller N400 than incongruous target words. In addition the evaluative decision on congruous target words was hypothesised to be faster and more accurate than on incongruous target words.

8.3.2 Methods

Participants

Fifteen musically trained (8 females) and 18 musically untrained (10 females) volunteers participated in the experiment. On average, musically trained participants were 25.4 years of age (sd 3.77) and musically untrained participants were 24.6 years of age (sd 4.01). Musicians had received approximately 14.6 years of musical training. All subjects were right-handed, native German speakers, with normal or corrected to normal vision and no hearing impairments.

Materials

The prime stimulus material consisted of 48 major chords, of which 24 were of a pleasant and 24 of an unpleasant musical timbre. The major chords were built either on the root (in C: c-e-g-c) or on the second inversion (in C: g-c-e-g) and played in each of the twelve keys. Chords were created in Cubase, Steinberg. Pleasant sounding chords were exported with the Grand option of Cubase and sounded like normal chords played on a piano. Unpleasant sounding chords were exported with the TinDrum option of Cubase and sounded considerably harsher and unpleasant. Chords were subsequently modified in Cool-Edit, possessing a final sampling rate of 44.1kHz and a 16-bit resolution. On average chords were 800 ms long. The choice of timbre initially depended on the judgement of the experimenter. A previous rating experiment (see Appendix A) conducted with both musically trained and untrained participants established that on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant, piano-timbre and tindrum-timbre chords were significantly different from one another in their perceived pleasantness (piano = 1.7 and tindrum = 4.3). There were no group differences in the pleasantness ratings.

To allow for a more fine-grained analysis, which parameters may be relevant for the perception of pleasantness in instrumental timbre, the presently used stimuli were analysed with regards to two parameters relevant for perceiving timbre: Attack time and spectral centroid. Attack time was calculated by extracting the Root-Mean-Square (RMS) of the signal over a time window of 10 ms, with a gliding window of 1 ms of the entire signal. RMS is a statistical measure of the magnitude of a varying quantity (in this case amplitude of a sound) and derived by means of the following formula: $\sqrt{[x]^2}$, where x denotes the arithmetic mean. The time from the beginning of the sound to the maximum RMS was calculated, which constituted the attack time. The spectral centroid measures the average frequency, weighted by amplitude of a spectrum. The standard formula for the average spectral centroid of a sound is: $c = \frac{\Sigma c_i}{i}$, where c_i is the centroid of one frame, and i is the number of frames for the sound. A spectral frame is the number of samples, which given the present stimuli was 44.1kHz. The frequency spectrum of each sound was calculated, by means of a Fast-Fourier-Transformation (FFT), with a size of 2048 sampling points to obtain an optimal estimate of spectral resolution given the present sampling rate of the

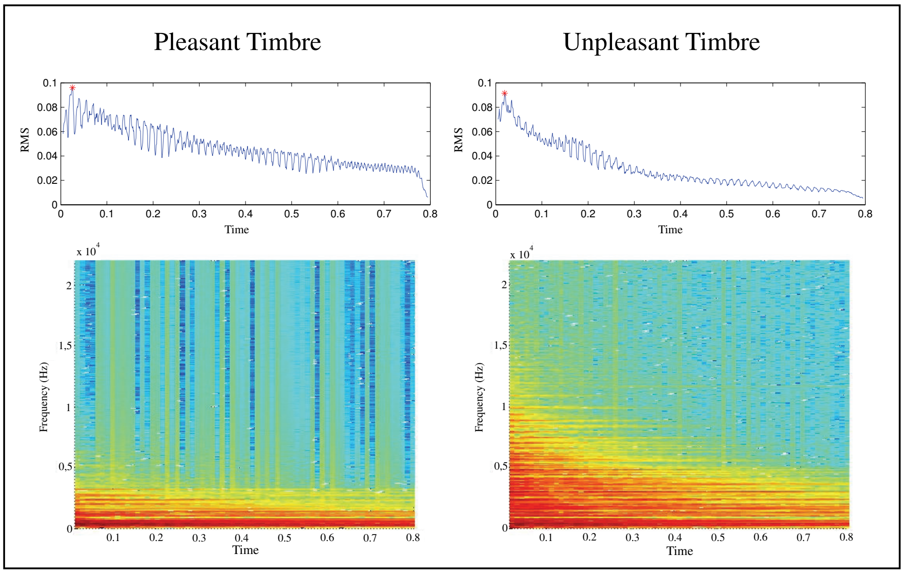

Figure 8.7: RMS (top) and Spectrogram (bottom) for an example of the pleasant (left) and the unpleasant (right) sounding chords (c-major in root position)

signal. Perceptually, the spectral centroid has been associated with the brightness of a sound (Schubert, Wolfe, & Tarnopolsky, 2004).

Attack time and spectral centroid were calculated for each chord, averaged for each timbre and compared using paired-samples t-tests. The Attack time was found to differ significantly between the pleasant (177 ms) and the unpleasant timbre (115 ms) $(t(23) = 4.85, p < .001)$. The spectral centroid was also found to differ between the two timbres and was considerably lower for the pleasant (402 Hz) than for the unpleasant timbre (768 Hz) $(t(23) = -10.79, p < .001)$. Chords with the unpleasant timbre appear to have a significantly earlier attack time as well as a brighter sound compared to chords with the pleasant timbre (see Figure 8.7).

Experimental target words, matching-, randomisation- and presentation procedures as well as ERP recording and data analysis were the same as in Experiment 1.

8.3.3 Results

8.3.3.1 Behavioural results

Reaction times and accuracy were analysed separately. The factors Prime and Target were entered into a repeated measures ANOVA in addition to the between-subject factor Training. Analysis of the reaction times revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between factors Prime and Target $(F(1,33) = 32.17, p < .0001)$, suggesting that both musically trained and untrained participants evaluated affectively congruous target words faster than affectively incongruous target words (see Figure 8.8). There were no further interactions or main effects.

The analysis of performance accuracy revealed high performance of both groups (musicians: 96.9%; nonmusicians: 94.9%). There were neither significant interactions nor any main effects, showing that error rates did not differ between groups nor were they sensitive to the relationship between valence of prime and target.

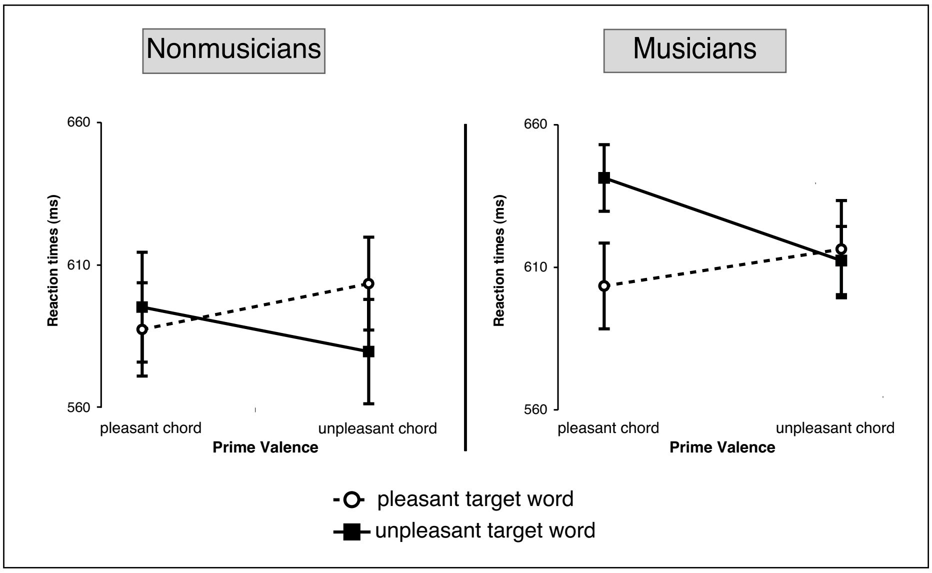

Figure 8.8: Mean reaction times (± 1 SE) for evaluative decisions on pleasant and unpleasant word targets

8.3.3.2 ERP results

Analysis of the ERPs in the time-window of 300-500 ms revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between factors Prime and Target $(F(1,33) = 17.88, p < .001)$, indicating a larger N400 for incongruous target words compared to congruous target words for both musically trained and musically untrained participants with a broad scalp distribution (see Figure 8.9). There were no further interactions with any of the other factors, nor any significant main effects.

In spite of suggestive visual evidence of further ERP effects in later time windows, these could not be statistically confirmed. For effects over more electrodes, see Figures C.5 and C.6 in Appendix C.

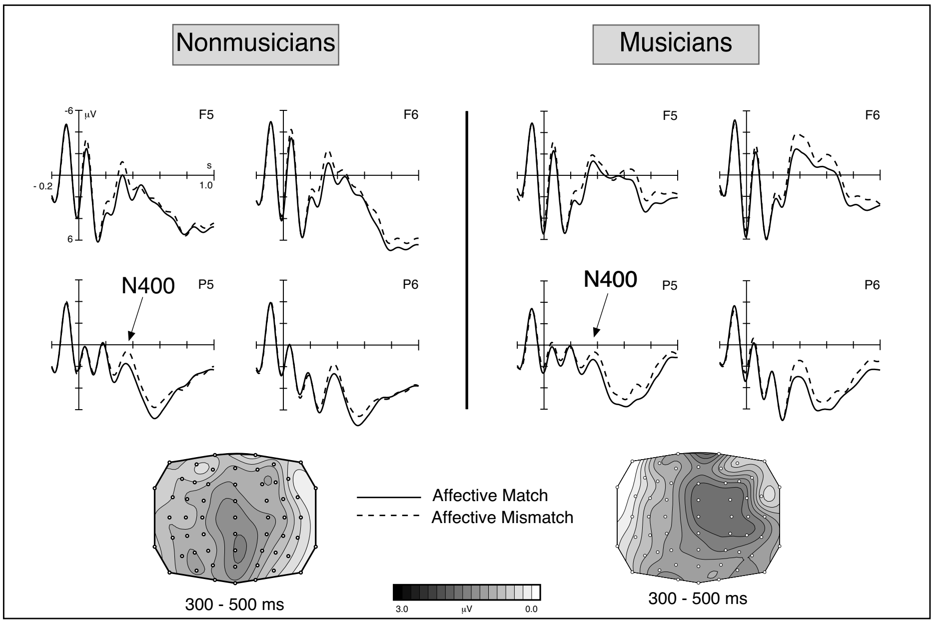

Figure 8.9: ERPs locked to onset of the target word. ERPs in response to an affective mismatch between prime and target valence (red line) resulted in a larger N400 between 300-500 ms compared to ERPs in response to a prime-target match (blue line). Both musicians and nonmusicians show the effect, which is broadly distributed over the scalp.

8.3.4 Discussion

This study demonstrates that timbre appears to be capable of communicating affective meaning. Both musically trained and untrained participants showed a significant behavioural priming effect, which in addition was accompanied by a larger N400 for target words mismatched in valence to the preceding prime. This provides support for a link between instrumental timbre and meaning in music, as has been hypothesised (Koelsch & Siebel, 2005).

Like in the first two experiments, no differences resulting from musical training could be found, neither in the behavioural data nor in the ERP data, suggesting once more that musical expertise does not modify the processing of affectively meaningful properties contained in basic features of the auditory input.

Similarly to the preceding two experiments, there was no priming effect found in the accuracy of responses. As the accuracy data shows, the task was very easy for participants and accuracy may therefore not have been sensitive to the prime-target relationship.

This is the first study to show that emotion perceived in music may depend on changes related to instrumental timbre. Whereas musical harmony, musical mode and tempo have been studied extensively, there is very little evidence of a link between timbre and emotion. This study shows that different instrumental timbres are perceived differently in terms of their valence. Seeing that the perception of instrumental timbre depends on more than one psychoacoustic property (Caclin et al., 2005), it cannot be said with any degree of certainty that any single one of those argued to be involved (e.g. attack time, spectral centroid) is uniquely responsible. Differences in timbre are perceived on the basis of the relative interplay of these psychacoustic properties and it is presumably the unique composition arising out of these, which make a timbre more or less pleasant sounding. Future work may want to focus more on the systematic exploration of each of the relevant (and further) psychoacoustic factors and to what extent these may be responsible for the perceived valence of instrumental timbre, an empirical issue for which the present paradigm seems very well suited.

Like in Experiment 1 and 2, the most parsimonious account for the presently observed affective priming effect and the associated N400 is spreading activation.

Equally, it is highly unlikely that the observed effect is a result of mere conflict at the response level resulting from the evaluative decision task used.

8.4 General Discussion of Experiments 1-3

In order to test the hypothesis that emotion in music is a primary means of communicating, several cross-modal priming experiments were conducted. Similar to semantic priming studies, musical chords were followed by target words which either matched or did not match the affect communicated by the chord. It was found that target words matching in valence to the preceding chord were evaluated faster than mismatched target words. In addition, ERP results showed an increased N400 in response to mismatched target words. The N400 has been seen as a classical index of semantic processing (Kutas & Federmeier, 2000) and therefore the present findings suggest that the emotion signalled by musical features is understood as meaningful, signalling different affective content. Thus emotions expressed by music appear to constitute a route to meaning in music.

The set of studies demonstrate the link between emotional expression in music and meaning using several musical parameters previously known to elicit emotions in listeners. Harmonic roughness, harmonic intervals and instrumental timbre are all an integral part of musical information, each of which are presumably processed and analysed separately (Koelsch & Siebel, 2005). It was shown that these individual aspects of the musical input are all capable of signalling affective meaning. Whereas this idea has already been put forward (Koelsch & Siebel, 2005), these studies are the first to provide any empirical evidence for this. The data speak for a link between musical features and meaning via emotional expression.

Whereas music is capable of expressing a whole range of emotions (e.g. fear, joy, anger; Krumhansl, 1997), the present set of studies investigated only affective categories such as pleasant/unpleasant and happy/sad. One may therefore view the present findings as the beginning, whereby affective catgories are shown to communicate affective meaning. Presumably this would not be different when employing a wider range of musical emotions, but this requires further empirical study.

The data do not speak on a link between the feeling of an emotion in response to and the perception of an emotion expressed by the music. No data were obtained which could inform on the emotional state of the participants. It seems plausible that the perception or recognition of an emotion expressed by music is sufficient to prime associated concepts and that this does not need to occur via the feeling of an emotion. However this entire issue is underresearched and requires more empirical attention to be able to give a definite answer.

The fast and accurate perception and recognition of an emotion would appear to be an important evolutionary strategy increasing the chances for survival of those who are best at it. In normal human exchange, music is not the primary means of communication, instead of which gestures, facial expressions and most importantly speech are used efficiently. It has been shown that the speech signal, apart from the use of words, contains important information signalling the emotional state of the speaker (Banse & Scherer, 1996; Schirmer & Kotz, 2006). It may be possible to conceive of similar mechanisms of perceptual analysis applied to both speech and music to scan for information signalling changes in the affective state. Even though the purpose of such mechanisms is lost on chords (as there is no defined signaller), it is not unlikely that mechanisms used for more relevant information (from an evolutionary perspective[9]) are also applied to music perception.

The present data give no indication of differences between musically trained and untrained participants with regards to the perception of affective meaning communicated by various musical features. Looking at both the behavioural responses and the ERP data, both groups seem to be influenced to the same degree by the affective congruity of chord-word pairs. This suggests that musical expertise has little effect on emotion and meaning related processes, which is in line with previous studies. Steinbeis et al. (2006) could show that there were no differences between musicians and nonmusicians in emotional responses to manipulations in harmonic exptancy violations and Bigand et al. (2005) demonstrated that there were no group differences in the perception of emotions in music.

In sum the present three experiments could show that the emotion expressed by various musical features (e.g. roughness, intervals, and timbre) is a pathway to meaning in music, which appears to be understood regardless of musical training and the recognition of which may be the result of the acoustic analysis of affect, which can be applied to speech and acoustic signals generally.

[9]This type of information would include all that is relevant for survival, including the emotional state of conspecifics.

Chapter 9

Experiment 4

Meaning and Tension-resolution patterns[10]

9.1 Introduction

The evidence so far in support of meaning in music includes extra-musical associations and sounds mimicking objects (Koelsch et al., 2004), and, as shown in the preceding chapter, emotions communicated by various musical features. Other routes to meaning in music, which have been hypothesised but for which there is no empirical evidence to date include tension-resolution patterns (Meyer, 1956; Swain, 1997). Such patterns do not refer to anything outside of the musical context and constitute a basic hallmark of all tonal compositions. This experiment is an empirical test of the hypothesis that these patterns represent a route to meaning in music.

Tension-resolution patterns constitute the relationship between musical elements (i.e. harmonies) based on their hierarchical organisation (i.e. Western major-minor tonal system). Studies have shown that the perceived closeness between two tones strongly depends on their belonging to the same key or not and also that within the same key, some tones will be perceived as more stable or final sounding than others (Krumhansl & Shepard, 1979). The same phenomenon has been described for the perception of chords (Krumhansl & Kessler, 1982). These perceptions mirror the harmonic distances as described by the Circle of Fifths and have been argued to follow a cognitive rule known as the *hierarchy of stability* (Bharucha & Krumhansl,

[10] This work was published in Steinbeis and Koelsch (in press).

1983). It has been demonstrated that this rule can be learnt implicitly through mere exposure to music in everyday listening situations (Tillmann et al., 2000).

Harmonic priming experiments have shown that listeners have expectations of subsequent harmonic progressions when hearing a single chord (Bharucha & Stoeckig, 1986). The degree to which listeners' expectations were fulfilled appeared to depend directly on the harmonic distance between prime and target chord pairs. Further studies have indicated that these expectations are related to the perception of musical tension (Bigand, Madurell, Tillmann, & Pineau, 1999; Bigand, Parncutt, & Lerdahl, 1996; Krumhansl, 1996; Krumhansl, 2002; Steinbeis et al., 2006), whereby listeners rate harmonic events as increasingly tense the further these occur away from the tonal root. Thus harmonic relationships are strongly linked with the perception of tension in music (Lerdahl & Jackendoff, 1983).

Meyer (1956) made an explicit link between tension-resolution patterns and meaning in music. He described meaning in music to arise out of its intrinsic structural properties, as opposed to extra-musical associations (as desribed by Koelsch et al., 2004). According to Meyer (1956), musical events do not symbolise anything (even though they can), but more importantly point to a musical consequence. Thus musical meaning arises out of the implicit suggestion of several possible subsequent events, constrained by the listeners expectations[11].

One way of finding out, whether tension-resolution patterns are meaningful to listeners is by studying ERPs. As already outlined, certain ERPs have come to reflect specific cognitive functions (e.g. the N400 for semantic processing, see Chapter 2). ERP studies investigating the processing of tension-resolution patterns have used harmonic expectancy violation paradigms in which a harmonic context is set up by means of a chord sequence generating expectations, which are then either confirmed or violated[12]. Several components have been reported, such as a P600 (Patel et al., 1998), an ERAN (Koelsch et al., 2000), as well as a later negativity distributed bilaterally over frontal electrodes known as the N500 (Koelsch et al., 2000). The findings of a P600 and an ERAN, components which are very similar to ERPs reported in the language literature on syntactic processing (Friederici, 2002), have led to the proposal that tension-resolution patterns are processed primarily in terms of structure and not content (Patel, 2003). This was confirmed by the absence of

[11]Presumably this is a fundamental mechanism by which meaning can arise in any context.

[12]A paradigm in essence similar to the way language processing is studied.

any N400-like response to harmonic expectancy violations (Besson & Macar, 1987; Janata, 1995). Whereas the P600 and the ERAN have been argued (Patel, 2003), and in the case of the ERAN clearly shown (Koelsch et al., 2005) to reflect the processing of structural violations, the role and nature of the N500 is not yet clear. Previous observations that the incremental decrease of the N500 to the increasing harmonic context (Koelsch et al., 2000) parallels the behaviour of the N400 to an increasing semantically plausible context (van Petten & Kutas, 1990), have suggested certain similarities. Thus, given the theoretical possibility that tension-resolution patterns may be meaningful to listeners and the presence of an ERP component of which the function is yet uncertain (N500), but which behaves similarly to another component known to reflect semantic processing (N400), this experiment set out to investigate whether tension-resolution patterns are meaningful by studying the functional specificity of the N500.

To test for the cognitive function of the N500, an interactive experimental paradigm was employed, using both music and language material (see Figure 9.1). This paradigm has already been successfully implemented (Koelsch et al., 2005) and allows to gage the relative influence of various types of language violations (syntactic or semantic) on the processing of harmonic expectancy violations (i.e. chord sequences which represent abstract entities of tension-resolution patterns). The fundamental assumption underlying the use of this paradigm was that of shared processing resources between the language and the music domain (as outlined in chapter 4)[13].

With regards to the present study, given that the N500 is reliably elicited in response to harmonic expectancy violations, it was hypothesised that if the N500 reflects semantic processing of harmonic expectancy violations, it ought to be reduced when occurring simultaneously with a semantic violation in the language material. Alternatively, if the N500 simply reflects more general working memory demands placed on the cognitive system as a result of any type of unexpected event, then a reduced N500 was hypothesised similarly for both types of language violations. Because the function of the ERAN has long been assumed to be of a syntactic nature

[13]This idea was put forward by Patel (2003) and argues that language and music share processing resources dedicated to structural aspects of the auditory input. These resources are limited and therefore processing a violation in language ought to reduce neural respources dedicated to processing a violation in music, if of similar functions (for details, see Chapter 4).

(for a review, see Koelsch, 2005), it was hypothesised that it would be modulated only by simultaneously presented syntactic language violations.

Several main effects for the language material were expected based on the literature, such as a LAN and a P600 for the syntactic gender violation (Gunter et al., 2000), as well as an N400 and possibly also a P600 for the low-cloze probability sentences (Kutas & Hillyard, 1984).

9.2 Methods

9.2.1 Participants

Participants were 26 right-handed musically untrained adults (13 males) with a mean age of 24.4 years, normal hearing and normal or corrected-to-normal vision.

9.2.2 Stimuli

Musical stimuli consisted of 78 five-chord sequences. Sequences were built in the following fashion in order to create maximal expectation towards the final chord: The first chord was the tonic of the ensuing sequence, the second chord was either a tonic, mediant, submediant or subdominant chord, the third chord was either a subdominant, dominant or dominant-six-four chord, the fourth chord was always a dominant-seventh chord, and the final chord was either a tonic chord or a Neapolitan chord. The tonic chord at fifth position represents the most expected chord whereas the Neapolitan chord represents a highly unexpected harmonic event. Tonic and Neapolitan chord occurred with equal probability (39 times in each language condition). The sequences were composed in different voicings and part writing occurred according to the classical rules of harmony. The chord sequences were generated with a piano sound (General MIDI No.1) under computerized control via MIDI on a Roland JV-2080 synthesizer (Hamamatsu, Japan). The first four chords were presented for 600 ms each, whereas the final chord was presented for 1200 ms. Sequences directly followed one another with no pause in between. Final chords had an approximate loudness of 55dB SPL, which was kept constant across all final chords. Each chord was played simultaneously with a single visually presented word. To ensure that participants would listen closely to the chord sequences, 30 chords deviating in timbre (harpsichord) were inserted pseudo-randomly into the auditory presentation. It was assumed that this would be sufficient to retrieve the N500 absent from the previous study (Koelsch et al., 2005). The timbre-deviants were not analysed further.

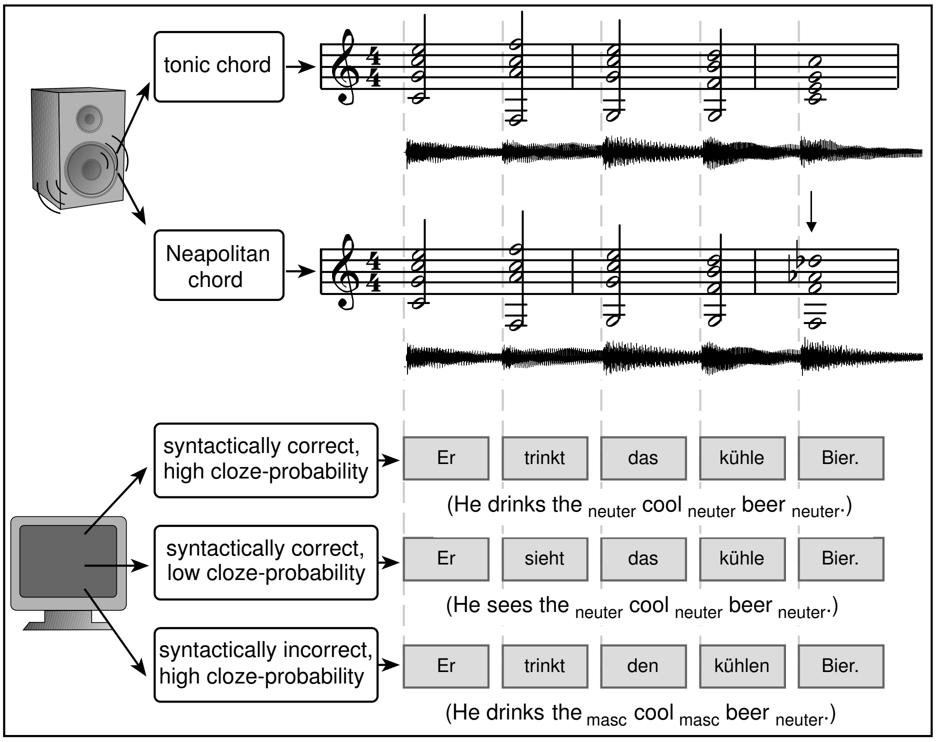

Figure 9.1: Sentences and musical material were combined in a way that each sentence type was presented with each harmonic sequence type creating a 3 x 2 design. Thus the following six experimental conditions were investigated: Sentences, which were both syntactically correct and semantically expected presented (1) with a regular or (2) with an irregular harmonic chord sequence; sentences, which were both syntactically incorrect and semantically expected presented (3) with a regular or (4) with an irregular harmonic chord sequence; sentences, which were both syntactically correct and semantically unexpected (low-cloze) presented (5) with a regular or (6) with an irregular harmonic chord sequence. It is important to note that the event indicating whether the presented material was regular or irregular always occurred at the end of the sequence (i.e. final word for sentences, final chord for chord sequences). Both sentences and chord sequences contained five elements, which were presented simultaneously. The final element represents the time-locked point of investigation to test the influence of sentence type on harmonic processing.

The same sentence material as the one used in previous studies (Gunter et al., 2000; Koelsch et al., 2005) was employed by the present study, which consisted of 39 sentences with five elements each. Each sentence was presented in three different versions, a syntactically correct and semantically expected one, a syntactically incorrect and semantically expected one and a syntactically correct and semantically unexpected one. The incorrect syntactic event consisted of a syntactic gender violation described in previous studies (Gunter et al., 2000), whereas the unexpected semantic event consisted of a low-cloze probability sentence completion at the final word, representing a very mild type of semantic expectancy violation. For the sake of simplicity, low-cloze probability sentences may be referred to as semantically unexpected or semantically incorrect, even though they were not semantically implausible. Each version of the sentence was presented with a tonic chord and a Neapolitan chord sequence (see Figure 9.1), producing 234 experimental sequences in total. The order of sentences was pseudo-randomised so that neither same sentence nor its altered version were presented in direct succession.

To ensure that participants paid attention to the language material, 16 memory questions were inserted pseudo-randomly, so that the experimental sequences were divided into 16 non-equally sized blocks.

9.2.3 Procedure

The experiment was conducted in an acoustically and electrically shielded EEG cabin. Subjects were seated in a comfortable chair facing a computer screen placed at approximately 1m distance. Participants were instructed to carry out two tasks. Firstly, they were told to detect chords played on a deviant instrument and signal this detection by pressing a button as soon as they heard the chord. This was done in order to retrieve the N500, which was absent in the study of Koelsch et al. (2005), presumably because attention was focussed exclusively on the language material. Additionally, participants were instructed to pay attention to the global meaning of each sentence, which they would be tested on infrequently during the experiment. This was done in order to ensure that the language material would also be sufficiently processed. Sixteen times a sentence appeared on the screen with the question underneath whether participants had already read the sentence before.

Eight of the sentences had been presented before, the remaining eight had not. The experimental session lasted approximately 28 minutes.

9.2.4 EEG recording and data analysis

The EEG recording and data analysis was the same as in Experiment 1-3.

For statistical evaluation, ERPs were analyzed by repeated-measures ANOVAs as univariate tests of hypotheses for within-subject effects. Mean ERP values were computed for four regions of interest (ROIs): left anterior (F7, F5, F3, FT7, FC5, FC3), right anterior (F8, F6, F4, FT8, FC6, FC4), left posterior (TP7, CP5, CP3, P7, P5, P3) and right posterior (TP8, CP6, CP4, P8, P6, P4). The factors entering the ANOVAs were the following: Chord type (expected [Tonic] and unexpected [Neapolitan]), Syntax (correct and incorrect), Cloze (high and low), Hemisphere (right and left) and AntPost (anterior and posterior). The time windows used for the analyses were 160-260 ms (ERAN), 600-800 ms (N500), 300-400 ms (LAN and N400) and 500-700 ms (P600). To test for a functional dissociation between ERAN and N500 at a later stage an additional factor of Time (160-260 ms and 600-800 ms) was used.

9.3 Results

9.3.1 Behavioural results

Participants detected 88.8 % of the timbre deviants and answered over 82.8 % of the memory questions correctly. This suggests that sufficient attention was directed to both information channels to process the input accurately.

9.3.2 ERP results

The presentation of the results is done in a hypothesis-driven manner. Because the study was carried out to observe effects of language material on music ERPs (and vice versa), the main effects are reported first and after that the relevant interactions.

Main effects: Language

Syntax The main effects which were calculated for ERPs in response to the syntactic gender violations are displayed in Figure 9.2 and reported in Table 9.1.

The syntactic gender violations elicited a LAN, which was followed by a globally distributed P600 (see Figure 9.2). For the time window of the LAN, there was a significant interaction between factors Syntax, Hemisphere and AntPost. A further ANOVA with factors Syntax and Hemisphere over anterior ROIs was also significant and the a priori hypothesis of an increased negativity over left anterior regions led to a significant effect for the factor Syntax. For the time window of the P600 there was a significant main effect of Syntax and no further interactions.

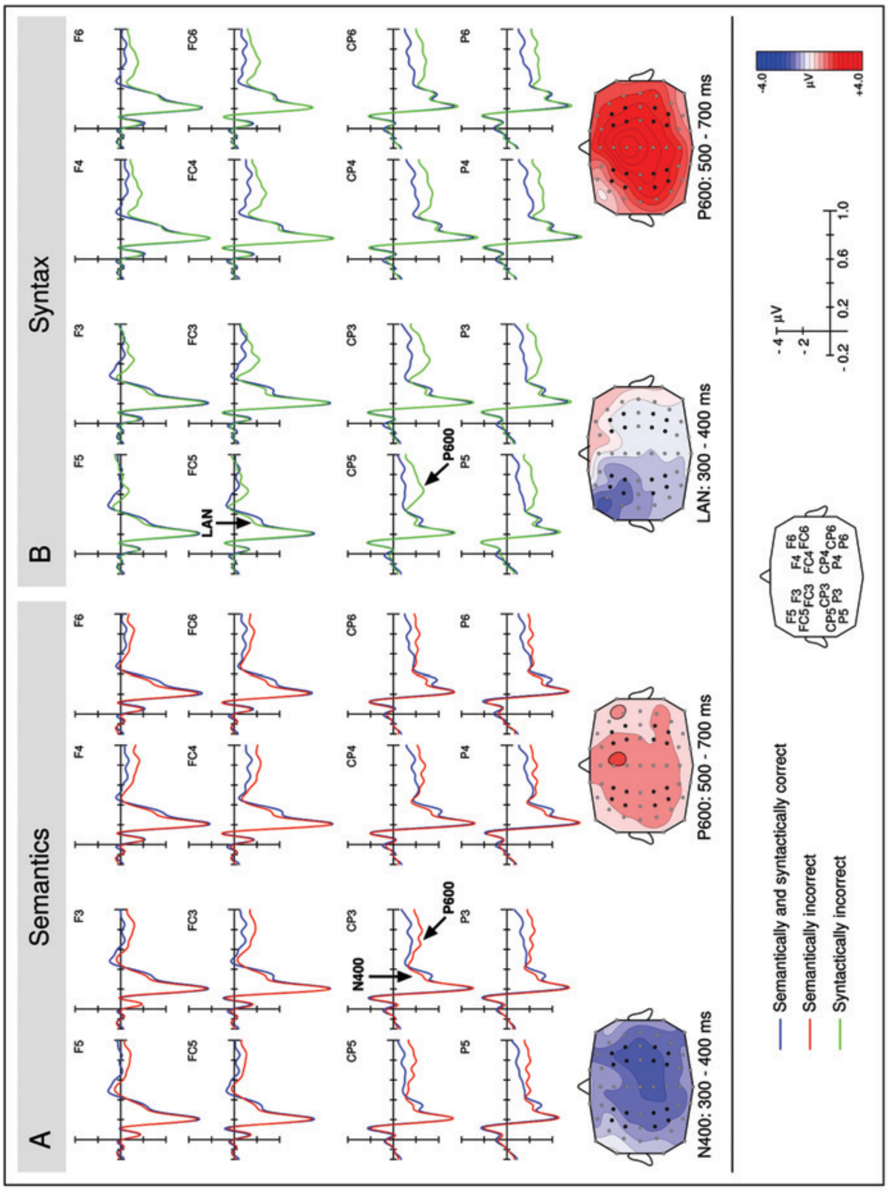

Figure 9.2: Low cloze probability sentences elicited an N400 with a central scalp distribution and a P600 (A); Syntactic gender violations elicited a LAN and a P600 (B)

Main Effects: Language Syntax				
LAN (300-400 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Syntax	1,25	1.09.	*ns*	
Syntax x Hem	1,25	5.58	< .05	3.63
Syntax x AntPost	1,25	6.84	< .05	4.72
Syntax x Hem x AntPost	1,25	10.78	< .01	0.33
LAN (300-400 ms) anterior ROIs				
Effect	*df*	*F value*	*p value*	*MSe*
Syntax	1,25	< 1		
Syntax x Hem	1,25	6.93	< .05	1.88
LAN (300-400 ms) left anterior ROI				
Effect	*df*	*F value*	*p value*	*MSe*
Syntax	1,25	5.46	< .05	3.84
P600 (500-700 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Syntax	1,25	21.17	< .0001	49.82
Syntax x Hem	1,25	< 1		
Syntax x AntPost	1,25	< 1		
Syntax x Hem x AntPost	1,25	< 1		

Table 9.1: Main effects of syntactic gender violations.

Semantics Main effects which were calculated for ERPs in response to the low cloze probability sentences are displayed in Figure 9.2 and reported in Table 9.2.

The low cloze probability sentences elicited an increased N400 compared to the high cloze probability sentences as was shown by an effect of Cloze between 300-400 ms. In addition, an increased P600 was found for the low cloze compared to the high cloze sentences as revealed by an effect of Cloze in the later time window (see Figure 9.2).

Main effects: Cloze probability				
N400 (300-400 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Cloze	1,25	7.24	< .05	12.27
Cloze x Hem	1,25	< 1		
Cloze x AntPost	1,25	< 1		
Cloze x Hem x AntPost	1,25	< 1		
P600 (500-700 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Cloze	1,25	8.83	< .01	15.01
Cloze x Hem	1,25	< 1		
Cloze x AntPost	1,25	< 1		
Cloze x Hem x AntPost	1,25	< 1		

Table 9.2: Main effects of cloze probability.

Main effects: Music

Main effects which were calculated for ERPs in response to the harmonic expectancy violations are displayed in Figure 9.3 and reported in Table 9.3.

ERAN The Neapolitan chord elicited a distinct ERAN maximal around 210 ms at right anterior sites (see Figure 9.3). An ANOVA for the time window 160-260 ms with the factors Chord Type, Hemisphere and AntPost revealed a significant three-way interaction indicating an increased negativity over right anterior sites than left posterior ones, a significant two-way interactions with the factors Chord Type and Hemisphere, indicating an increased negativity over the right hemisphere than over the left, a significant two-way interaction with the factors Chord Type and AntPost, indicating an increased negativity over anterior sites than over posterior ones, as well as a significant effect of Chord Type.

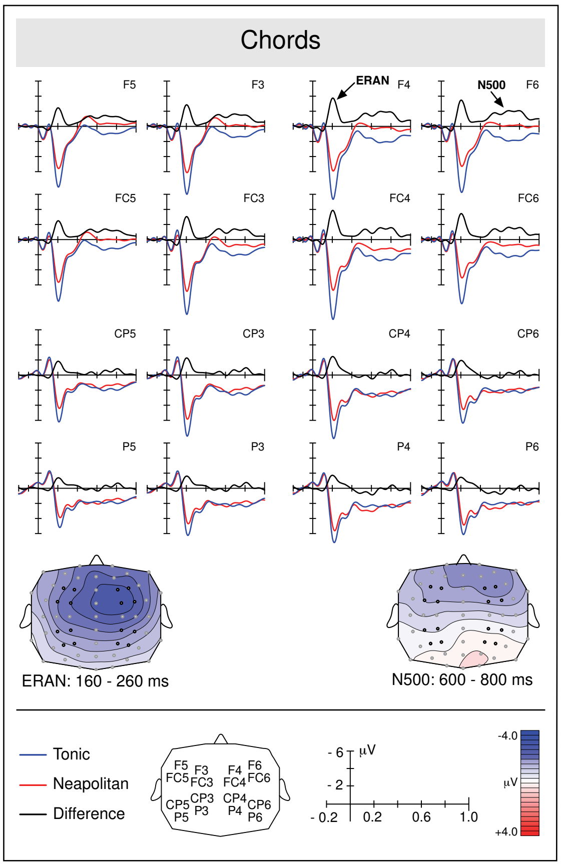

Figure 9.3: Harmonically unexpected chords elicited an ERAN and an N500 with a bilateral frontal scalp distribution

N500 The Neapolitan Chord also elicited a clear N500, with an onset at around 450 ms and lasting up to around 900 ms and a maximum at 650 ms, with an anterior distribution and a slight right-hemispheric weighting (see Figure 9.3). For statistical analysis however, the time window of 600-800 ms was opted for, which was motivated by an inspection of interactions with the language material in this particular latency band. As a result the main effect of Chord Type is also reported in this time window over all sentence types.

An ANOVA for the time window 600-800 ms with the factors Chord Type, Hemisphere and AntPost revealed a significant three-way interaction, indicating an increased negativity over right anterior sites compared to left posterior ones, a significant two-way interaction with the factors Chord Type and AntPost, indicating an increased negativity over anterior sites compared to posterior ones and a significant main effect of Chord Type.

Main effects: Harmonic expectation				
ERAN (160-260 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type	1,25	98.14	$< .0001$	294.9
Chord Type x Hem	1,25	9.81	$< .01$	4.43
Chord Type x AntPost	1,25	39.01	$< .0001$	17.76
Chord Type x Hem x AntPost	1,25	10.29	$< .01$	2.25
N500 (600-800 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type	1,25	8.52	$< .01$	21.73
Chord Type x Hem	1,25	< 1		
Chord Type x AntPost	1,25	30.01	$< .0001$	25.1
Chord Type x Hem x AntPost	1,25	9.01	$< .01$	1.78

Table 9.3: Main effects of harmonic expectation.

Interactions between language violations and music - ERAN and N500

Interactions between ERPs in response to harmonic expectancy violations and the different types of sentence violations are displayed in Figures 9.4 and 9.5 and reported in Table 9.4.

Syntax and ERAN When the Neapolitan Chord was presented simultaneously with the syntactic gender violation, the ERAN amplitude was reduced compared to when presented with syntactically correct sentences. An ANOVA for the time-window 160-260 ms with the factors Chord Type and Syntax revealed a significant interaction (see panel A in Figure 9.4).

Syntax and N500 When the Neapolitan Chord was presented simultaneously with the syntactic gender violations, there was no significant reduction of the N500.

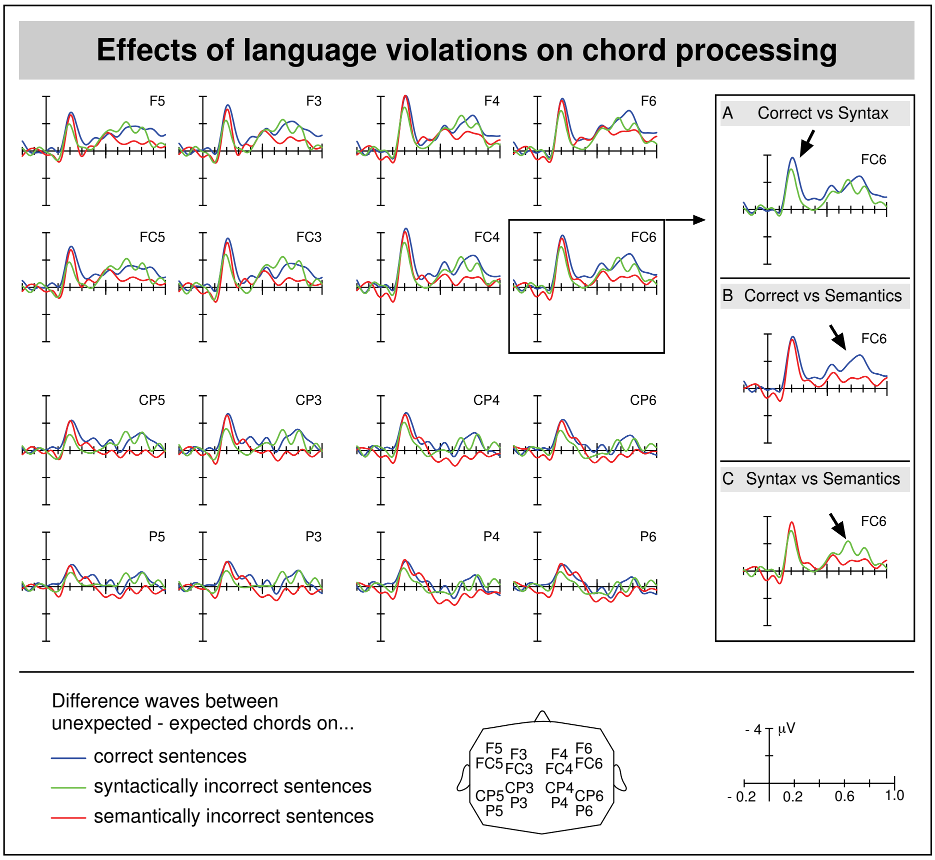

Figure 9.4: Sentences with a syntactic gender violation reduced the ERAN but not the N500 (A); semantically unexpected sentences reduced the N500 but not the ERAN (B); semantically unexpected sentences reduced the N500 more than syntactic gender violations (C), providing the strongest evidence for semantic specificity of the N500.

Semantics and ERAN When the Neapolitan Chord was presented simultaneously with the semantically unexpected sentences, there was no significant reduction of the ERAN, as confirmed in an ANOVA with the factors Chord Type and Cloze.

Semantics and N500 There was a clear interaction between factors Chord Type and Cloze in the N500 time window, showing a reduced N500 when the Neapolitan was presented with semantically unexpected sentences compared to when presented with semantically expected sentences (see panel B in Figure 9.4).

Interactions: Music ERPs with language material				
ERAN (160-260 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type x Syntax	1,25	6.44	< .05	16.73
Chord Type x Syntax x Hem	1,25	< 1		
Chord Type x Syntax x AntPost	1,25	< 1		
Chord Type x Syntax x Hem x AntPost	1,25	< 1		
Chord Type x Cloze	1,25	< 1		
Chord Type x Cloze x Hem	1,25	< 1		
Chord Type x Cloze x AntPost	1,25	< 1		
Chord Type x Cloze x Hem x AntPost	1,25	< 1		
N500 (600-800 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type x Syntax	1,25	< 1		
Chord Type x Syntax x Hem	1,25	< 1		
Chord Type x Syntax x AntPost	1,25	< 1		
Chord Type x Syntax x Hem x AntPost	1,25	< 1		
Chord Type x Cloze	1,25	8.26	< .01	19.68
Chord Type x Cloze x Hem	1,25	< 1		
Chord Type x Cloze x AntPost	1,25	< 1		
Chord Type x Cloze x Hem x AntPost	1,25	< 1		

Table 9.4: Interactions of music-related potentials with language conditions.

Functional specificity of ERAN and N500 This section reports statistical effects for specific modulations of the ERAN and the N500. These are derived from interactions of several factors, which have all been reported above.

Results from ANOVAs to test for the specific modulation of music-related potentials with the different types of language violations are displayed in Figure 9.5 and reported in Table 9.5.

To test for a functional dissociation between the two components, an ANOVA was conducted using the additional factor Time (time window ERAN/time window N500), Chord Type and Language Violation (Syntax and Cloze) as factors. A significant three-way interaction was found reflecting that the ERAN was modulated more by the incorrect language syntax than by the unexpected language semantics and that the opposite was the case for the N500 (see Figure 9.5).

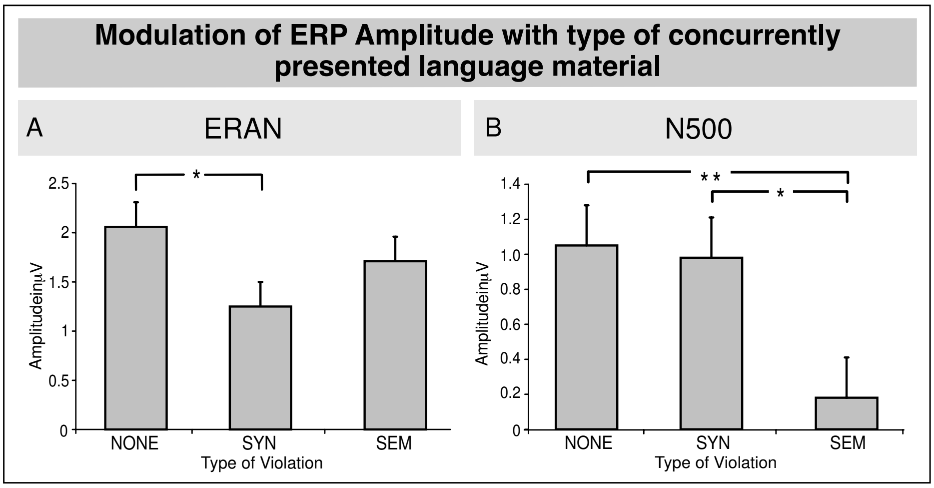

Figure 9.5: The ERAN was significantly reduced when presented concurrently with a syntactic violation compared to when presented with correct sentences, but not semantically unexpected sentences (A); The N500 was significantly reduced when presented concurrently with a semantically unexpected sentence, compared to both correct sentences and syntactically incorrect sentences (B).

An additional test for a functional specificity of both the ERAN and the N500 was used to compare their responsiveness to either type of language violation. This test was performed with an additional ANOVA over the time windows of both the early (160-260 ms) and the late (600-800 ms) component using type of language violation (Syntax and Cloze) as an additional factor. Conducting an ANOVA with factors Chord Type and language violation over the ERAN time window, no interaction was found, suggesting that the ERAN was not significantly more reduced for the syntactic violation than for the semantically unexpected sentences. Running the same type of ANOVA over the later time window however, it showed that the N500 was significantly smaller when the Neapolitan chord occurred concurrently with the semantically unexpected sentences than the syntactic gender violations.

Functional specificity of ERAN and N500				
Dissociation of ERAN and N500				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type x Syntax x Cloze x Time	1,25	7.41	< .05	20.43
Specific modulation of ERAN				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type x Syntax x Cloze	1,25	1.67	< .2	
Specific modulation of N500				
Effect	*df*	*F value*	*p value*	*MSe*
Chord Type x Syntax x Cloze	1,25	5.73	< .05	16.53

Table 9.5: Functional specificity of ERAN and N500.

Interactions between music violation and language - LAN and N400

The fact that this experiment employed a dual-task compelled to look at the effect of the harmonic expectancy on the language ERPs. Interactions between ERPs in response to the sentence violations and harmonic expectancy are reported in Table 9.6.

LAN The LAN was reduced when presented concurrently with the Neapolitan chord compared to when presented with the tonic chord (see Figure 9.6). An ANOVA with factors Chord Type and Syntax over the time window 300-400 ms revealed a significant two-way interaction, replicating the results reported by Koelsch et al. (2005).

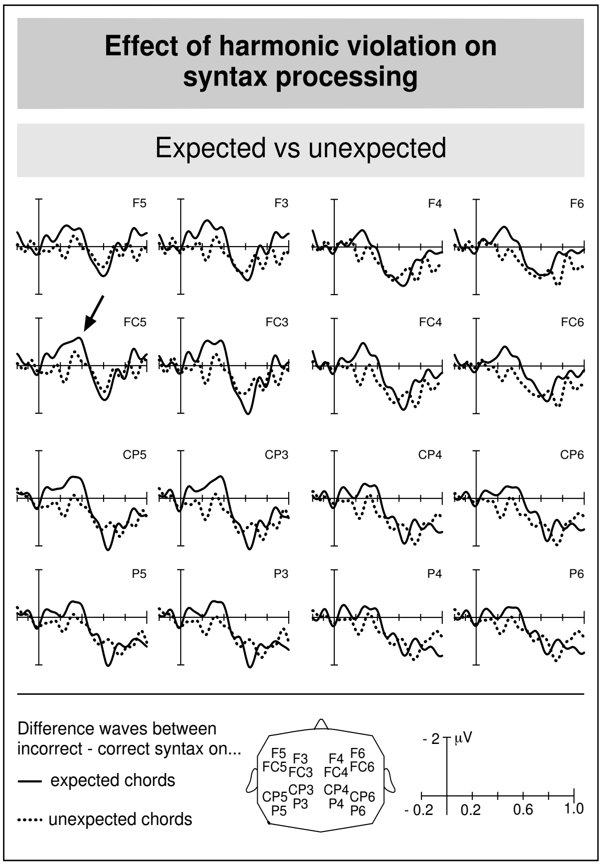

Figure 9.6: The LAN was significantly reduced when presented concurrently with a harmonically unexpected chord compared to when presented with a harmonically expected chord

Interactions: LAN with harmonic expectation				
LAN (300-400 ms)				
Effect	*df*	*F value*	*p value*	*MSe*
Syntax x Chord Type	1,25	6.44	< .05	16.73
Syntax x Chord Type x Hem	1,25	< 1		
Syntax x Chord Type x AntPost	1,25	< 1		
Syntax x Chord Type x Hem x AntPost	1,25	< 1		

Table 9.6: Interactions of syntactic gender violations with harmonic expectation.

N400 No reduction of the N400 could be found when presented simultaneously with the Neapolitan chord. This also replicates the previous results of Koelsch et al. (2005) and is possibly more surprising in this case, as a clear N500 was elicited unlike in the previous study. This however, suggests that whereas the syntactic interaction (LAN-ERAN) works both ways, the semantic interaction (N400-N500) only works in the direction of music.

9.4 Discussion

This study reports several significant main effects all in accordance with the literature. The LAN and P600 elicited by the syntactic violations and the N400 and P600 elicited by the semantically unexpected sentences demonstrate that the sentence material was processed sufficiently to allow for an observation of its effect on the music material. This is confirmed by the relatively high accuracy performance in the memory test. The harmonic expectancy violations elicited both a clear ERAN and an N500, their time-course and scalp distribution conforming to previous results (Koelsch et al., 2000). The fact that the N500 was at all present can be attributed to task requirements, showing that it was possible to retrieve the N500, a primary component of interest.

The main finding of this study however is that the N500 was modulated only by the semantically unexpected sentences and not the syntactic gender violations. Because the semantically and the syntactically anomalous material can be considered to make comparable demands on attentional and working memory resources and the fact that the N400, the semantic component, and the LAN, the syntactic component occur in identical time windows, the two anomalous sentence types represent optimal control conditions with regards to more general cognitive mechanisms. Also, this effect cannot be interpreted in terms of overlapping scalp topographies between the P600 elicited for the low-cloze probability sentences and the N500 in the music condition, as the syntactic violation also elicited a P600 and should therefore have reduced the N500, which it did not. Thus, the fact that there was a significant interaction between sentence violation and N500 amplitude, suggests that this modulation is specific to semantic processing and cannot be attributed to more general attentional or working memory demands. This is therefore a clear piece of evidence that language and music appear to share processing resources with regards to building meaning, which in turn implies for the present data that meaning can arise out of music which does not refer to anything outside itself, confirming musicological and recent scientific speculations (Meyer, 1956; Koelsch et al., 2004).

The finding of an interaction between the processing of harmonic structure and semantic language material suggests that structure in music can convey meaning and enables one to make sense of the music. Two possible mechanisms might be able to account for this. One is similar to accounts of how meaning can arise merely out of structure in language (Ullman, 2001). For instance a sentence without any propositional meaning such as “Clementia glicked the plag” (example taken from Ullman, 2001) still communicates the basic information that Clementia performed an action on an object based on knowledge of word class and rules of hierarchical composition. Thus the type of musical meaning arising out of harmonic structure may be subserved by a mechanism dedicated to processing structure more generally, establishing meaningful relationships between the compositional elements. The other possible mechanism is that of mediation via affect. It has been shown that harmonic expectancy violations give rise to emotional responses (Steinbeis et al., 2006). As already demonstrated in the previous studies, emotional communication appears to convey meaning in music, as a result of which one may argue that the present findings indicate an additional affective route to meaning via musical structure. Further work is still required to be able to choose between these two potentially rivalling accounts.

This study shows that music and language also share neural resources on a semantic level. Recent reviews on the neuroscience of language processing (Friederici, 2002; Bookheimer, 2002) haved argued for a "dual-system" of semantic processing, with procedures, such as contextual integration on the one hand, and the representation of semantic knowledge on the other. Because the N500 has been shown to reflect a sensitivity to the build-up of harmonic context (Koelsch et al., 2000), much like the N400 reflects the establishment of a semantic context (van Petten & Kutas, 1990), it is likely that the reduction of the N500 with semantically unexpected material results from the shared neural processes dedicated to building a semantic context. Recent reviews on semantic language processing and the function of the N400 (Kutas & Federmeier, 2000) have highlighted contextual integration as a crucial contributor to the generation of this component, making the possibility of an overlap between music and language on a level of contextual integration plausible.

One could therefore argue that the shared resources are dedicated to semantic procedures (i.e. context building), which echoes theoretical accounts of the nature of shared resources dedicated to syntactic structure building between music and language (Patel, 2003). One may speculate on the involvement of the pars orbitalis (BA47) of the inferior frontal gyrus to be involved in such operations. Language studies have systematically shown a role of BA47 in semantic processing (for a review, see Bookheimer, 2002), whereas recent studies have shown increased activations of this region when processing musical structure, compared to auditory input without any coherent temporal structure (Levitin et al., 2003).

This study provides evidence for one of several pathways in music capable of generating meaning. Presumably there are differences in the underlying mechanisms of the kind of pathway reported here and other pathways reported in previous studies (Koelsch et al., 2004). As described above, it is most likely that tension-resolution patterns are meaningful in that they represent instances of highly constrained context-building mechanisms. The musical stimuli used by Koelsch et al. (2004), however seem to tap into semantic representations of music more directly, by means of imagery or association. Therefore, it appears as if the multitude of pathways to meaning in music may be subserved by different mechanisms, which are indicated by differing ERP components, the N400 and the N500.

The additional finding of a specific interaction between the ERAN and the syntactic language violations only, confirms a considerable string of previous results surmising the ERAN to reflect the processing of musical syntax (Koelsch et al., 2005; Koelsch & Siebel, 2005; Koelsch, 2005). The fact that the ERAN was only reduced when presented concurrently with the syntactic language violation and not the semantically unexpected sentences suggests that the ERAN is modulated by the recruitment of syntactic processing resources required also by the language system. However, seeing that there was no significant interaction of the ERAN amplitude with the type of language violation (i.e. not smaller with the syntactic than the semantic violation), implies that this reduction may also have something to do with increased working memory operations placed on the system, to which the ERAN may be sensitive.

The observation that the latency of the LAN (300-400 ms) is later than that of the ERAN (160-260 ms) and yet the latter was reduced when reading material that typically elicits a LAN, is not simple to interpret. Possibly other neural processes not detected by scalp electrodes might play a role in such an interaction.

The fact that the same pattern as reported in a previous study (Koelsch et al., 2005) can be observed, whereby the LAN is significantly reduced by concurrently presented harmonic irregularities provides further evidence that the ERAN may reflect syntactic processing and lends further support to the hypothesis that music and language share syntactic integration resources (Patel, 2003; Koelsch et al., 2005).

The finding that musical and linguistic syntax mutually affect one another, whereas only linguistic semantics affects musical semantics and not the other way round confirms previous studies (Koelsch et al., 2005). This presumably reflects the fact that the semantic features of language are considerably more salient than those of harmonic irregularities. Whereas there are so far no studies to support this assumption, one might predict a two-sided interaction when using more semantically salient musical material, such as the one used in the study by (Koelsch et al., 2004). Behavioural evidence to the contrary (Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005) could be interpreted in terms of differing task-requirements and stimulus material. A more complete investigation taking varying attention to the concurrently presented stimulus material and task requirements into account might be a better candidate for responding to this particular issue. Future research might also want to consider the possibility of presenting both language violations within the same sentence material, and to observe the effects of music violations on both the LAN and N400.

This study investigated whether a basic characteristic of Western tonal musical compositions can be meaningful to music listeners. The finding of shared neural resources involved in semantic processing between language and musical structure suggests that this is indeed the case. This implies that in some way all musical pieces can be meaningful to listeners familiar with the basic structural properties of music (Meyer, 1956). This meaning can be mediated via feelings or emotions (Meyer, 1956) or may arise out of a knowledge of musical rules and regularities, which can be implicitly learned through exposure to music (Tillmann et al., 2000). This can be taken as evidence for intrinsic musical properties and formal musical structures to be understood as meaningful by listeners and that meaning in music can occur without any reference or association to other meaningful events, objects, or sounds.

Chapter 10

Experiments 5 and 6

Comparing the neural representations of meaning in language and music

10.1 Experiment 5

10.1.1 Introduction

To consider the nature of meaning in music one is invariably compelled to think about the way meaning is communicated in language. Obviously, whereas the mechanisms for each are presumably quite different, and in the case of music much more circumscribed, specific and less flexible, an additional question arises with regards to the representation of meaning of both music and language. As shown in the previous experiment, there is some evidence that meaning in music and language share processing resources, which in turn suggests comparable brain structures to be involved in processing meaning in language and music. This study used a paradigm very similar to the one employed in the first three experiments, but changed the order of prime and target, in order to see whether incongruous chord targets can also elicit an N400, therefore indicating similarities in the representation of meaning in language. A brief introduction to previous attempts at eliciting an N400 for music is given below.

The quest for an N400 in response to musical targets has a long history. Taken as the classic marker for semantic processing (Kutas & Hillyard, 1980) several studies aimed to see whether an N400 can be elicited by contexts not explicitly semantic (Besson & Macar, 1987; Janata, 1995; Patel et al., 1998). As already shown in the discussion on ERPs to musical targets (see Chapter 3), no N400 has been reported in response to musical context violations, as a result of which it was concluded that these musical contexts (and therefore music in general) do not convey meaningful information.

Recent studies have shown that music can convey semantically meaningful information under certain circumstances (Koelsch et al., 2004). Rather than using musical targets, Koelsch et al. (2004) used music to prime concepts and words to probe semantic memory. Words are of course efficient probes, because they pack a lot of information into a single perceptual entitiy. Deservedly, Janata (2004) raised the pertinent question whether an N400 could be elicited by a musical piece priming needle, when followed by another piece priming river. Seeing that the analysis of ERPs always requires a precise onset from which to average the signal it is difficult to answer this question. Presumably, if one were to conduct a gating experiment with the aim to discover the precise point at which the meaning of a piece of music is recognised one would be a step closer.

Another way of adressing the question whether a single musical element is capable of eliciting an N400 (and therefore efficiently probing semantic memory) is by changing the context in such a way that a single target can communicate enough to be considered meaningful.

Experiment 1-3 have already demonstrated that a single musical element is enough to communicate affective meaning. Thus, if these contain sufficient information, then theoretically it ought to be possible to use them as targets to probe for the N400. The present experiment made use of a slightly modified version of the affective priming paradigm used in Experiments 1-3, with the difference that this time affective primes were pleasant and unpleasant words and targets were consonant and dissonant chords (see Figure 10.1).

The reason harmonic roughness was opted for as initial test of eliciting an N400, was that perceptually these chords are highly saliently different and given the considerable literature on affective responses covarying with harmonic roughness, this would represent a suitable starting point to adress this issue. Thus as a result of the recent study on meaning in music and returning to a question that has long interested cognitive scientists interested in music, the present study aimed to adress, whether, when using an appropriate context, musical targets can elicit an N400. It way hypothesised that should musical targets be able to elicit an N400, targets incongruous with the affect of the preceding word prime ought to do so more than when congruous with the preceding affective word. In addition, relatedness between target and prime should also affect the response times and possibly also the accuracy with which the target is evaluated, where congruent target chords should elicit faster and more correct responses than incongruent target chords.

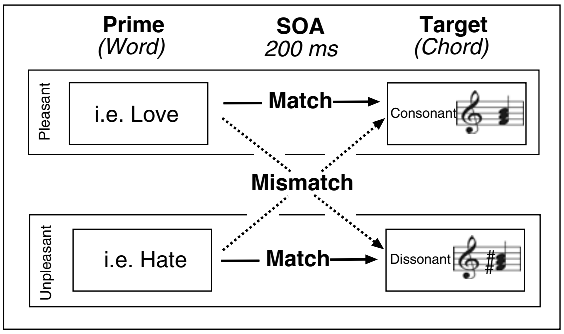

Figure 10.1: Design of the affective priming paradigm. In order to see whether musical targets can elicit an N400 words are used as affctive primes and chords as targets

Given the lack of differences between musically trained and musically untrained participants in the first three experiments, no specific hypothesis was formulated with regards to musical training effects and an interaction with prime-target affect. However, given that musicians learn, practice and play music for several hours a day, it is likely that musicians are generally faster at evaluating musical targets than nonmusicians.

10.1.2 Methods

10.1.2.1 Participants

Twenty musically trained (10 females) and 20 musically untrained (10 females) volunteers participated in the experiment. On average, musically trained participants were 26.2 years of age (sd 1.22) and musically untrained participants were 25.8 years of age (sd 1.51). Musicians had received approximately 14.9 years of musical training. All subjects were right-handed, native German speakers, with normal or corrected to normal vision and no hearing impairments.

10.1.2.2 Materials

The materials were identical to the ones used in Experiment 1. For each word, one pleasant and one unpleasant target chord were chosen, which was done randomly and altered for each participant. Thus each word was presented twice, followed once by a congruous chord and once by an incongruous chord. The experimental conditions and number of trials were the same as in Experiment 1.

10.1.2.3 Procedure

The procedure was almost identical to the one used in Experiment 1, with the exception of reversing prime and target (chord primes and word targets to word primes and chord targets; see Figure 10.1) Words appeared on the screen for 200 ms after which a chord was presented from two loudspeakers positioned to the left and right of the participant. Participants were instructed to decide as fast as possible whether the chord sounded pleasant or unpleasant. Seeing that the perception of pleasantness in music is a subjective matter, accuracy of responses was not emphasised. Responses were made with a button-box, pressing left for pleasant and right for unpleasant, which was switched after the first half of the experiment.

10.1.2.4 EEG recording and data analysis

Recording and analysis were the same as in Experiment 1, apart from the time windows chosen for statistical analysis. Based on visual inspection of the data, the chosen time-windows ranged from 200-400 ms and 500-700 ms.

10.1.3 Results

10.1.3.1 Behavioural results

Reaction times and accuracy were analysed separately. The factors Prime and Target were entered into a repeated measures ANOVA in addition to the between-subject factor Training. Analysis of the reaction times revealed a significant effect of group ($F(1,38) = 20.22, p < .0001$) indicating that musically trained subjects evaluated the targets significantly faster than musically untrained subjects. There was also a three-way interaction between the factors Prime, Target and Training ($F(1,38) = 9.71, p < .005$), suggesting further differences resulting from musical training. A separate analysis of the data for musically trained and untrained subjects showed a significant two-way interaction between factors Prime and Target only for the musically trained group ($F(1,18) = 26.771, p < .001$) and none for the musically untrained group ($p > .5$). Thus, only musically trained participants evaluated affectively congruous target words faster than affectively incongruous target words (see Figure 10.2).

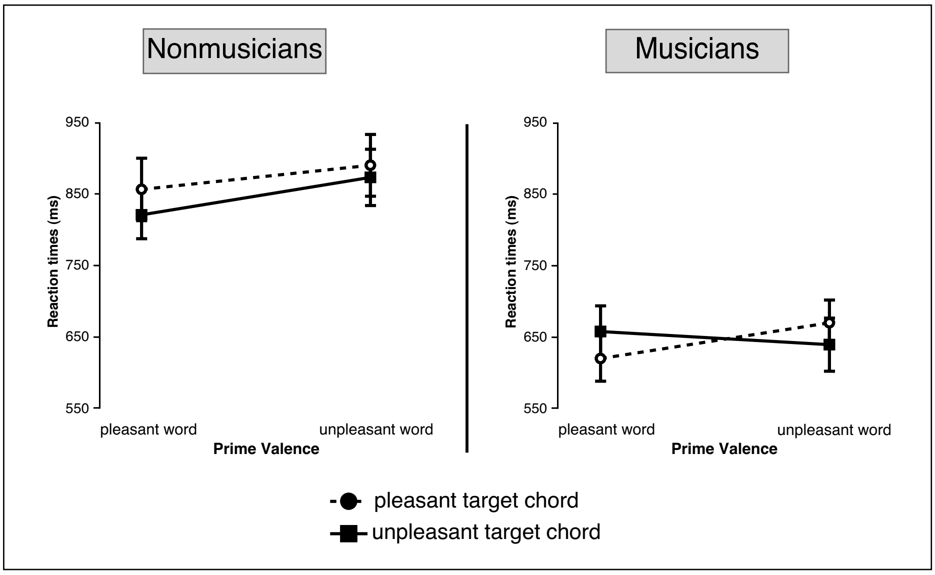

Figure 10.2: Mean reaction times (± 1 SE) for evaluative decisions on pleasant and unpleasant chord targets

To check that the group differences in the behavioural priming effect were not simply due to response speed, reaction times of the ten slowest musicians and the ten fastest nonmusicians were compared and analysed. The two groups' response times were no longer statistically different, whereas the difference in the interaction with the factors Prime and Target persisted ($F(1,19) = 10.21, p < .005$). The interaction between Prime and Target was still significant for the ten slowest musically trained responders ($F(1,9) = 26.771, p < .001$) but not so for the ten fastest musically untrained responders.

Tendentially nonmusicians evaluated chords more slowly after unpleasant primes than after pleasant primes, but this effect did not reach significance. There were no further interactions or main effects, apart from a basic effect of group.

The analysis of performance accuracy revealed relatively high performance of both groups (musicians: 90.8%; nonmusicians: 88.2%). This shows that a high percentage of the chords' expressed emotion was recognised accurately. There were neither significant interactions nor any main effects, showing that error rates did not differ between groups nor were they sensitive to the relationship between valence of prime and target.

10.1.3.2 ERP results

Visual inspection showed a negativity between 200-400 ms for the musically trained group, which was almost completely absent in the musically untrained group (see Figure 10.3). Analysis of the ERPs in the time-window of 200-400 ms revealed a significant three-way interaction between the factors Prime, Target and Training $(F(1,38) = 5.08, p < .05)$. A separate analysis of the data from musically trained and untrained subjects showed a significant two-way interaction between factors Prime and Target only for the musically trained group $(F(1,19) = 20.23, p < .001)$ and none for the musically untrained group $(p > .8)$, indicating a larger negativity for incongruous target chords compared to congruous target chords, but only for musically trained participants. In this time window, there were no further interactions nor main effects.

In addition to the negativity, both groups showed a positive component with a centro-parietal maximum between 500-700 ms. Analysis of the ERPs in the time-window of 500-700 ms revealed no significant three-way interaction between the factors Prime, Target and Training, but a significant two-way interaction between the factors Prime and Target $(F(1,38) = 7.04, p < .05)$. There was a further interaction of factors Prime, Target and AntPost with the factor Training $(F(1,38) = 12.01, p < .005)$, suggesting the interaction between Prime and Target to be larger for musically untrained subjects than for musically trained subjects over anterior electrodes $(F(1,38) = 9.72, p < .005)$, but not over posterior electrodes $(p > .3)$. There was also an interaction between the factors Prime, Target and AntPost $(F(1,38) = 7.04, p < .05)$ and a separate analysis over anterior and posterior ROIs showed a significant interaction between factors Prime and Target only over posterior electrodes ($F(1,38) = 15.43, p < .001$).

In spite of the suggestive visual evidence for an additional later and frontally distributed negativity for the musically trained group, this could not be confirmed statistically. For effects over more electrodes, see Figures C.7 and C.8 in Appendix C.

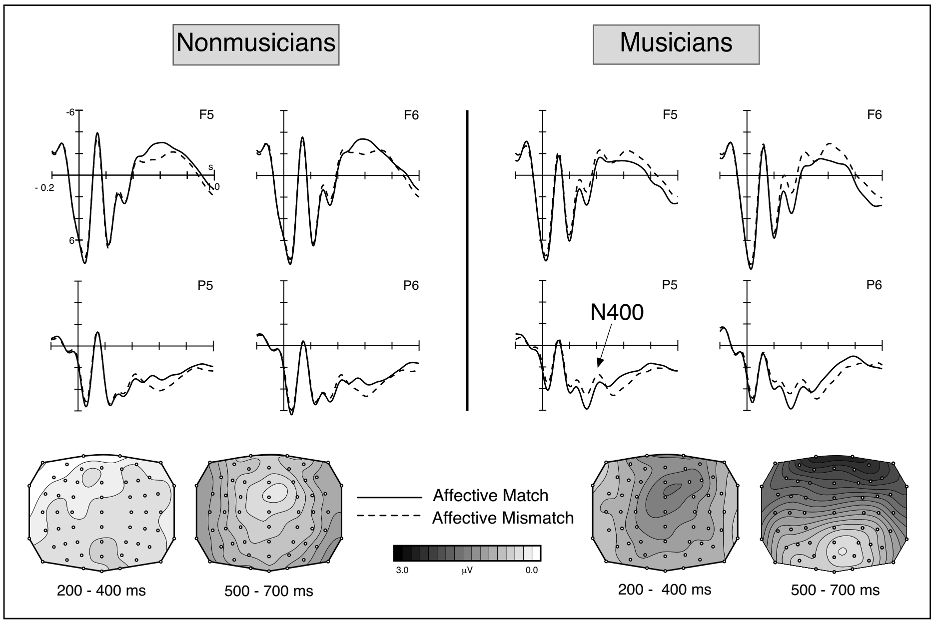

Figure 10.3: ERPs locked to onset of the target chord. ERPs in response to an affective mismatch between prime and target valence (red line) resulted in a larger N400 between 200-400 ms compared to ERPs in response to a prime-target match (blue line) for musicians only. An additional later positivity between 500-700 ms is displayed by both groups.

10.1.4 Discussion

This experiment investigated whether meaning in music is represented in a comparable fashion to the meaning expressed by language, by attempting to elicit an N400 for single musical targets. Having shown in the first three experiments that single chords can communicate affective meaning, a set of these chords was now used as targets congruous or incongruous with the preceding affective context. The behavioural data indicate that whereas musicians generally responded faster than nonmusicians, they were also the only group which demonstrated a behavioural priming effect, where congruous chord targets were evaluated faster than incongruous chord targets. This difference persisted even when there were no group differences anymore in the overall response speed (as shown in the sub-group comparison), which suggests that the observed group differences are not a mere artefact of response speed but due to other factors, which will be explained in conjunction with the ERP data below.

Like in the first three experiments, there was no priming effect visible in the accuracy scores. The scores however indicate that the task was relatively easy and ceiling effects may therefore have precluded seeing any effects in the accuracy of responses.

Analysis of the ERPs showed similarities as well as differences between the two groups. Both groups displayed a late positive component, which, due to the absence of a behavioural priming effect for the musically untrained group, cannot be associated with the behavioural effect in the musicians. In spite of differences in scalp distribution between the two groups, it is assumed that the component reflects the same cognitive mechanism. It may be that this positivity reflects the general processing of musically unexpected events, as has been hypothesised to be the case for melodic and harmonic deviants (Besson & Faita, 1995; Regnault et al., 2001). The main group difference in the ERP data however is that only the musically trained group showed a negativity in a time window roughly approximating that of the N400.

The present finding of a clear behavioural priming effect and the presence of a negativity for musically trained subjects suggests that this effect may genuinely reflect an N400. However, as in the first three experiments, the alternative explanation of the behavioural effect merely representing a mismatch at the response level and the associated ERP mirroring this conflict (as an N200), as opposed to the hypothesised semantic mismatch cannot be entirely ruled out. The rivalling account of an N200 receives some support from the observation that in the present experiment, both the time course and the distribution roughly match that of the N200 reported previously (van Veen & Carter, 2002). What may speak against this interpretation is the issue why the groups should differ with regards to such a simple mechanism. The response conflict of incongruous trials is present for both groups, therefore the question arises why only the musically trained group should show a component reflecting this conflict. This inconsistency between the groups with regards to such basic mechanisms sheds doubt on the N200 interpretation.

However, the only way that these two rivalling accounts of the present effect (N200 and response conflict *vs* N400 and semantic mismatch) can be teased apart would be by trying to specifiy the source of the ERP. Whereas the N200 has been argued to be generated in areas known to subserve response conflict, such as the ACC (van Veen & Carter, 2002), the N400 has been argued to be generated in areas such as the inferior frontal gyrus as well as the middle temporal gyrus (Maess et al., 2006; Koelsch et al., 2004). Thus this issue cannot be entirely resolved and discussed until further analyses or studies have been conducted. As a result, the next experiment used fMRI, employing the same paradigm as in the present study with musically trained subjects only. Because the exact nature of the group difference depends on the level of mismatch at which musicians process the incongruous chords, and which in turn can only be told once the fMRI data has been obtained, the discussion of group differences with regards to the present effect will be postponed and discussed jointly with the fMRI data.

10.2 Experiment 6

10.2.1 Introduction

This study was motivated by two questions. For one, the nature of representations of meaning in music and language can also be studied with regards to brain areas, specifically whether there may be a neural locus dedicated to processing meaning generally including that communicated by music as well as language. Even though source localisations conducted by Koelsch et al. (2004) have suggested an involvement of the middle-temporal gyrus (MTG), specifically BA 21/37 bilaterally in processing semantic relationships between pieces of music and target words, there is still no evidence on a direct comparison between the integration of meaning in music and in language.

The neural correlates of semantic processing in language have been intensely studied. As the extensive review in chapter 2 suggests, semantic processing has been suggested to be subserved by left inferior frontal regions, notably the pars opercularis (BA 47; for a review, see Bookheimer, 2002; Friederici et al., 2003; Schirmer et al., 2004), as well as the middle temporal gyrus (Kiehl et al., 2002; Kuperberg et al., 2000; Newman et al., 2001; Rodd et al., 2005; Rissman et al., 2003). Presumably, the neural generators of the N400 can also be located in those regions, which has been confirmed by recent source localizations (Maess et al., 2006). In addition to the studies on semantic language processing, the study by Koelsch et al. (2004) suggests that the bilateral MTG is involved in processing semantic relationships between music and language. Also, a recent study using a very similar paradigm to the one employed in Experiment 1 (Schirmer et al., 2004), reported activations of both left and right inferior frontal gyrus (IFG), interpreting the left IFG as indicator for processing semantic aspects of emotional prosody, and right IFG (BA 44/45) as processing the conflict at the response level[14]. To see if these structures were also involved in the processing of affective meaning as primed by single chords, the same paradigm was used as in Experiment 1. It is therefore hypothesised that when using chords as primes and words as targets to probe semantic memory, significant activations in the medial temporal lobe (BA 21/37) and in the left IFG can be observed.

[14]The study by Schirmer et al. (2004) also used an evaluative response task producing conflict at the response level.

An additional motivation were the findings reported in Experiment 5, which have highlighted the need to provide further evidence in order to decide between two rivalling accounts of the negativity found for incongruous chord targets compared to congruous ones for the musically trained group. Explaining the effect at the level of a basic stimulus-response conflict or at the level of an affective/semantic conflict generates differing predictions with regards to the neuroanatomical areas involved. Thus the second part of the study will be used to provide answers on the conflicting accounts of the negativity found for incongruous chord tagets by using the same experimental paradigm as in Experiment 5.

Conflicts at the stimulus-response level have been shown to be reflected by an N200. Recent studies have shown that the source of the N200 lies in the cACC (van Veen & Carter, 2002) and that response conflict also activates this region as shown in an fMRI study (van Veen, Cohen, Botvinick, Stenger, & Carter, 2001). Therefore, the ACC seems to be critically involved in exerting control at the response level. In addition to the ACC the involvement of the right IFG has also been reported when exerting control at the response level (Milham et al., 2001) an interpretation which was also used to account for right IFG activation in the study of Schirmer et al. (2004). Thus, it is hypothesised that if the negativity in response to incongruous targets reported for musically trained subjects in Experiment 5 reflects response conflict processing, incongruous chord target trials should activate the cACC more strongly than congruous chord trials as well as the right IFG.

As shown above, the neural correlates of semantic processing are distinct from those reported for response conflict processing, the former comprising left inferior frontal and medial temporal lobe structures. Thus, if the effect reported in Experiment 5 is an N400, reflecting semantic processing, incongruous chord target trials should activate the left IFG and medial temporal structures more strongly than congruous chord trials.

The rationale of the present design assumes that incongruous targets will recruit brain structures which are involved in that type of processing more strongly than congruous targets. To adress the two issues raised above, the experimental paradigms of Experiment 1 and Experiment 5 were used and combined in a single fMRI study. Seeing that the behavioural priming effect and the negative ERP were only obtained in the musically trained group, the present experiment was conducted only with musicians.

10.2.2 Methods

10.2.2.1 Participants

Sixteen musically trained (8 females) volunteers participated in the experiment. On average, participants were 24.7 years of age (sd 4.36) and had received approximately 15.6 years of musical training. All subjects were right-handed, native German speakers, with normal or corrected to normal vision and no hearing impairments.

10.2.2.2 Materials

The materials used were identical to the ones used in Experiments 1 and 5.

10.2.2.3 Procedure

This experiment used both the paradigm of Experiment 1 and the paradigm of Experiment 5. The presentation was identical as outlined under the respective experiments. To avoid a serial effect, eight subjects started with a block of chord primes and word targets (henceforth Block Chordprime) and the remaining eight started with a block of word primes and chord targets (henceforth Block Wordprime). Thus, since each experimental paradigm included 96 trials, the scanning session comprised 192 experimental trials, as well as 20 silence trials (or null events), which were used as baseline condition. There were therefore 212 trials in total. In order to obtain measures of the hemodynamic response at various points, the presentation of successive trials varied between 6.0 to 8.0 seconds in 0.4 s steps. Each participant underwent a practice session outside of the scanner until they could perform each task with 90% accuracy. In spite of this extensive training, one participant still managed to confuse the task during the experiment and was excluded from the subsequent analysis. The scanning started with 10 dummy trials, so that participants could accustom to performing the task in the scanning environment.

10.2.2.4 fMRI acquistion and analysis

Imaging was performed on a 3T Trio scanner (Siemens, Erlangen, Germany) equipped with the standard bird-cage head coil. Stabilization cushions were used to reduce head motion. Two sets of two-dimensional anatomical images were acquired for each participant immediately prior to functional imaging. A T_1 Model Driven Equilibrium Fourier Transform (MDEFT) sequence (data matrix = 256 x 256; recycle time (TR) = 1.3 s, time echo (TE) = 10 ms) and an EPI-T_1 sequence were used. The EPI-T_1 had the same parameters as the sequence with which the fMRI data was acquired. A total of 22 axial slices (19.2 cm field of view (FOV), 64 x 64 matix; 5 mm slice thickness, 1 mm spacing) parallel to the AC-PC plane were acquired using a single-shot, gradient recalled EPI-sequence (TR 2000 ms, TE 30 ms, 90° flip angle). Nine-hundred-and-ninety time-point functional runs with each time point sampling the 22 slices were recorded.

The fMRI data processing was performed using the software package LIPSIA (Lohmann et al., 2001). The functional data were corrected for motion artefacts offline with the Siemens motion correction protocol (Siemens). To correct for the temporal offset between slices acquired in one scan, a cubic-spline interpolation was applied. A temporal highpass filter with a cutoff frequency of 1/128 Hz was used for baseline correction and a spatial Gaussian filter with 3.768 mm full width at half maximum was applied. To align the functional slices with a 3D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (three rotational, three translational) was performed. The rotational and translational parameters were acquired as a result of the MDEFT and EPI-T_1 slices, to achieve an optimal match between the slices and the indiviual 3D reference data set. This 3D data set had been acquired for each subject individually at a previous scanning session. The MDEFT volume data set with 160 slices and 1 mm slice thickness was standardised to the Talairach stereotaxic space (Talairach & Tournoux, 1988). The rotational and translational parameters were subsequently transformed by linear scaling to a standard size, so that the resulting parameters could be used to transform the functional slices by using trilinear interpolation so that the functional slices were aligned with the stereotaxic coordinate system.

Statistical evaluation was based on a least-squares estimation using the general linear model (GLM) for serially autocorrelated observations (Worsley & Friston, 1995; Friston et al., 1995). The design matrix was generated using a synthetic hemodynamic response function, and its first and second derivative (Friston et al., 1995). The model equation, including the observed data, the design matrix, and the error term, was convolved with a Gaussian kernel, with a dispersion of 4 s full width at half maximum. Contrast images of the differences between the specified conditions were calculated for each subject. The individual contrast images were then entered into a second-level random effects analysis (one sample t-test). Subsequently t-scores were transformed into Z scores. To protect against false-positive activations, only regions with a Z score > 3.09 and with a volume > 135mm^3 (5 voxels) were considered and subsequently reported.

Because specific regions of interest (ROI) were specified a priori in the experimental hypotheses, specific ROI analyses were also conducted over inferior frontal and medial temporal structures. ROIs were defined based on the local maxima in a contrast map. In order to reduce smearing of activity from different anatomical structures, the local maximum was considered if it was within a radius of 15 mm. Significant maxima were used as the center of spherical ROIs, which had a radius of 8 mm. To compare condition specific activations, subjectwise contrasts between the critical conditions were calculated, for which normalised Z scores were averaged over all voxels within a ROI. The resulting mean Z scores were then subjected to an ANOVA with the factors Prime and Target as repeated measures factors.

The GLM is not as suitable to estimate whether particular brain structures are *not* involved as the basis of specific functions, as other statistical procedures such as Bayesian statistics. Instead of testing the estimated probability to obtain a contrast activation with the null hypothesis of no activation being true (i.e. p values), the Bayesian approach directly infers the probability that a contrast between two experimental conditions is larger than zero (Neumann & Lohmann, 2003). When applied to second-level analyses, the posterior probability of the group's contrast to be larger than zero is calcuated based on the parameter estimations for the individual participants on the first level. The resulting posterior probability maps provide a measure of the reliability of activation differences in the group (expressed as a percentage value between 0 to 100). This approach was used to both check for the presence as well as the absence of activations in particular cortical regions.

10.2.3 Results

10.2.3.1 Behavioural results

Block Chordprime Reaction times and accuracy were analysed separately. The factors Prime and Target were entered into a repeated measures ANOVA. Analysis of the reaction times revealed a significant interaction between the factors Prime and Target ($F(1,14) = 23.12, p < .001$), showing that incongruous target words were evaluated faster than congruous target words (see Figure 10.4). There were no main effects of neither Target nor Prime.

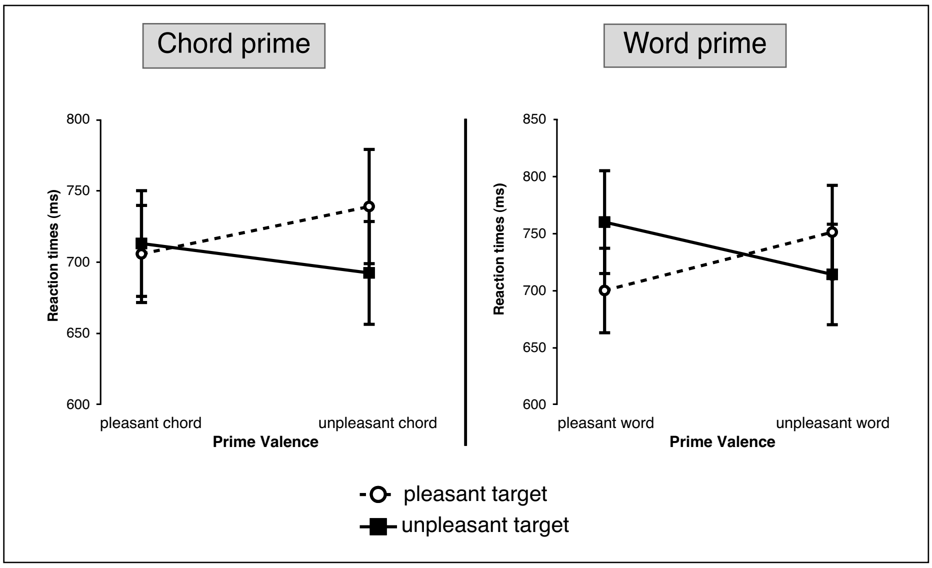

Figure 10.4: Mean reaction times (± 1 SE) for evaluative decisions on pleasant and unpleasant word targets on the right and for pleasant and unpleasant chord targets on the left

Block Wordprime Reaction times and accuracy were analysed separately. The factors Prime and Target were entered into a repeated measures ANOVA. Analysis of the reaction times revealed a significant interaction between the factors Prime and Target ($F(1,14) = 21.32, p < .001$), showing that incongruous target chords were evaluated faster than congruous target chords, replicating the behavioural effect reported in Experiment 5 (see Figure 10.4). There were no main effects of neither Target nor Prime.

The analysis of performance accuracy revealed a high performance for both Block Chordprime: 94.4%; as well as for Block Wordprime: 98.9%. The analysis of performance accuracy revealed neither significant interactions for the two blocks nor any main effects for the two blocks.

10.2.3.2 fMRI results

Block Chordprime In order to determine which brain areas are specifically involved in the processing of affective meaning of music, as primed by music and probed by a word, all incongruous trials in the block Chordprime were contrasted against all congruous trials in the block Chordprime. It was found that the left anterior insula and the right middle temporal gyrus were preferentially engaged (see Table 10.1 and Figure 10.5).

In order to see whether these activations showed the required interaction between the valence of chord prime and word target, ROI analyses were conducted over the local maxima of the two regions for each condition and subjected to a repeated measures ANOVA. A significant interaction between Prime and Target was found for both the left anterior insula ($F(1,14) = 20.47, p < .001$) as well as the right MTG ($F(1,14) = 14.9, p < .005$). There were neither main effects of Prime nor main effects of Target for any of the structures reported above.

Block Chordprime: *incongruous > congruous trials*							
Region	x	y	z	BA	Peak Z-value	Extent (mm^2)	Bayes %
L anterior insula	-38	15	0	-	3.44	162	99.95
R MTG	61	-24	-9	21/37	3.70	270	99.95

Table 10.1: Activations for target words when contrasting incongruous trials against congruous trials at a threshold of p < 0.001 (uncorrected) and a minimal cluster size of 5 voxels. The last column indicates posterior probabilities from Bayesian statistics estimating the reliability of activation differences.

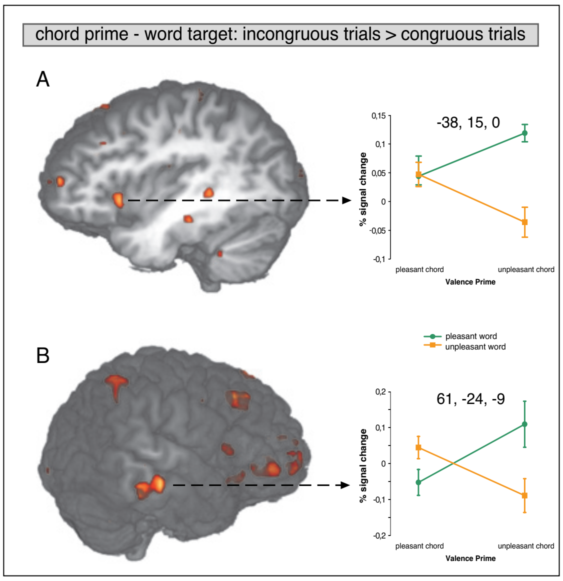

Figure 10.5: Increased activations in the left anterior insula (A) and the right MTG (B) when contrasting incongruous trials against congruous trials, displayed at an uncorrected threshold of $p < 0.005$. Other cortical regions which may look active did not survive the applied threshold.

Block Wordprime To assess which brain areas were involved in the increased processing of chord targets incongruous to the preceding affective context set by the word primes and if these structures are more typically engaged in response conflict or semantic processing, all incongruous trials in the block Wordprime were contrasted against all congruous trials in the block Wordprime.

It was found that the right IFG and the right superior temporal sulcus (STS) were preferentially engaged (see Table 10.2 and Figure 10.6). There was no activation of the ACC even at a very liberal threshold of $p < .1$ uncorrected. Using the Bayesian probability maps it was shown that the caudal ACC was activated with a probability of 40%. Whereas this is already quite low, this activation can be accounted for by "spillage" from surrounding cortical regions. To see whether the activations of rIFG and rSTS show the required interaction between the valence of word prime and chord target, ROI analyses were conducted over the local maxima of the two regions for each condition and subjected to a repeated measures ANOVA.

Block Wordprime: *incongruous > congruous trials*							
Region	x	y	z	BA	Peak *Z*-value	Extent (mm^2)	Bayes %
R IFG	43	36	0	47	3.62	162	99.95
R STS	40	-45	3	-	3.32	216	99.95

Table 10.2: Activations for target chords when contrasting incongruous trials against congruous trials at a threshold of $p < 0.001$ (uncorrected) and a minimal cluster size of 5 voxels. The last column indicates posterior probabilities from Bayesian statistics estimating the reliability of activation differences.

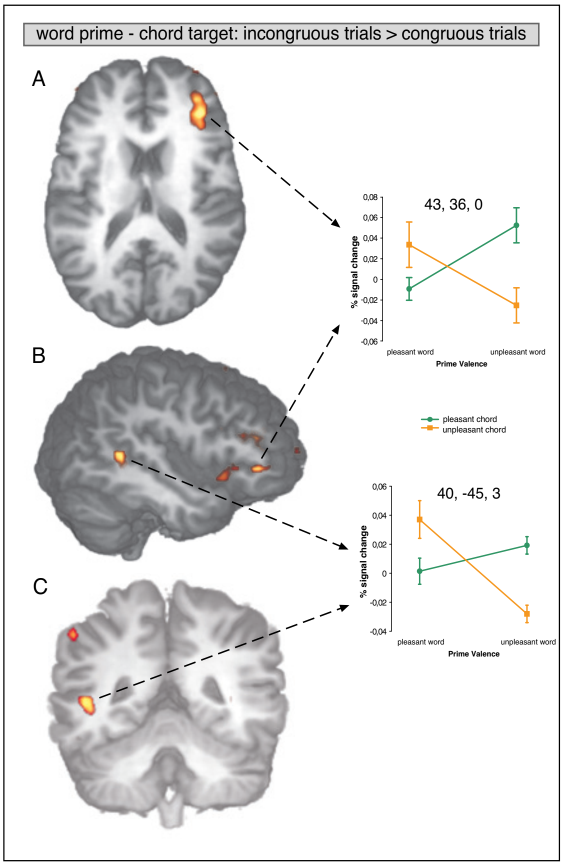

Figure 10.6: Increased activations in the right IFG (panels A & B) and the right STS (panels B & C) when contrasting incongruous trials against congruous trials, displayed at an uncorrected threshold of $p < 0.005$. Other cortical regions which may look active did not survive the applied threshold.

A significant interaction between Prime and Target was found for both the right IFG ($F(1,14) = 15.2, p < .005$) as well as the right STS ($F(1,14) = 21.73, p < .001$). There were neither main effects of Prime nor main effects of Target for any of the structures reported above.

10.2.4 Discussion

The study adressed whether processing meaning in language and in music is subserved by the same neural structures, thus indicating comparable representation in the brain. It also served to uncover the neural correlates of musical target processing specifically, in order to account for data from Experiment 5. The present discussion will deal with the neural correlates of word target and chord target processing in turn, after which the correlates for each will be compared.

Using single chords as primes and words as targets, the first part of the study aimed to see whether target words affectively incongruous with the preceding musical context would activate structures which have been implicated in studies on semantic processing using a similar paradigm (affective priming) but different material (prosodic speech; Schirmer et al., 2004) or using a different paradigm (semantic priming) but similar material (music; Koelsch et al., 2004). The behavioural data revealed a significant affective priming effect replicating the data from Experiment 1, whereby target words matched in valence to the preceding prime were evaluated significantly faster than mismatched target words.

When contrasting incongruous trials against congruous trials significant activations were observed in the left anterior insula and the right MTG. The left anterior insula activation also encompassed the left IFG at lower thresholds, as reported by Schirmer et al. (2004), which did not however survive a more conservative threshold. Whereas the left IFG has been associated with task-related or strategtic mechanisms of semantic priming (Friederici, 2002), the left anterior insula has already been reported for semantic processing (Friederici et al., 2003) and connected with automatic processing (Mummery, Shallice, & Price, 1999). This would be in line with the use of the present SOA of 200 ms, which presumably only taps into automatic rather than strategic priming processes. More importantly though, the anterior insula seems to play a crucial role in the evaluation of emotional stimuli

(Calder, Lawrence, & Young, 2001; Phillips et al., 1997), which is plausible given the evaluative decision task used.

In addition to anterior insula activations, the present data also suggest an involvement of the right MTG (BA 21/37). The activation found in the present experiment for incongruous word targets is close to the coordinates reported for the source localisation of the N400 in the paper by Koelsch et al. (2004, x = 44.75; y = -36.95; y = -2.65). It has recently been argued that the MTG represents an important structure in the so-called ventral processing stream of language comprehension, which hosts a lexical interface mapping sounds to meaning (Hickok & Poeppel, 2000, 2007). Even though the word targets were presented visually and thus did not *sound* at all, it is plausible that the posterior MTG maps the meaning of all language input, regardless of modality. This study thus provides further evidence that single chords are capable of conveying meaningful information, which when incongruous to other meaningful stimuli, recruits brain structures known to process the semantic properties of incoming information.

It is highly likely that the activations found in the present group of musically trained subjects would also extend to musically untrained subjects, since Experiment 1 already indicated a high degree of comparibility between the two groups in terms of the behavioural response to the affective priming using chords as primes, as well as the associated electrophysiological response.

The second part of the study was to provide unequivocal evidence for either of two possible explanations used to account for the findings in Experiment 5 and if so, to see whether the correlates are comparable to those found for processing language meaning. Using words as primes and chords as targets the aim was to see if activations are found in areas involved in the regulation of response conflict (i.e. ACC and rIFG) or rather in areas known from semantic processing (medial temporal lobe).

The behavioural data revealed a significant affective priming effect replicating the data from Experiment 5, whereby target chords matched in valence to the preceding prime were evaluated significantly faster than mismatched target chords. When contrasting incongruous against congruous chord targets there was no significant activation of the ACC and the Bayesian maps indicated a 60% probability that this was not the case. The right IFG however was activated, suggesting that the conflict at the response level was still processed, but did not recruit other structures typically involved (i.e. ACC). This is in line with previous findings (Schirmer et al., 2004). Studies typically reporting ACC activations usually employ Stroop paradigms with material that is exclusively connected to a motoric response and not the processing of other stimulus features (affective properties; Milham et al., 2001; van Veen & Carter, 2002), which may explain the absence of ACC activation in the present study. This implies that the ACC is not involved in processing affective incongruity. Seeing that this structure has been shown to be the generator of the N200 (van Veen & Carter, 2002) and that using the same paradigm as in Experiment 5 does not activate the ACC, the ERP reported cannot be an N200. The alternative explanation is that of an N400, suggesting that the emotion communicated by a single chord incongruent with its context is represented in a similar manner to words semantically incongruous with their preceding context.

The interpretation that the ERP reported in Experiment 5 constitutes an N400 receives further support from the activation of the right STS reported in the present study. The present activation of the STS is only slightly superior to the activation of the MTG in the Chordprime Block and is an area that is still considered to be part of the ventral processing stream dedicated to mapping sounds to their meaning (Hickok & Poeppel, 2000, 2007). It has been argued that the specific role of the posterior STS is at the level of phonological processing (Liebenthal, Binder, Spitzer, Possing, & Medler, 2005; Hickok & Poeppel, 2000, 2007). With regards to the present data, this interpretation is implausible, since chords do not possess phonological information. However, seeing that the posterior STS appears to discriminate between potentially meaningful (speech) and meaningless (nonspeech) information, this suggests a functional role in processing meaning on a more basic level. Once the information is successfully identifyied as speech (or coherent language), it is relayed to the posterior MTG for lexico-semantic processing. Seeing that chords do not posssess a lexical entry, they were not processed further in the posterior MTG, explaining the difference to the target word activation.

Thus, it appears that single chords when incongruent with the preceding affective context activate structures which are usually involved in identifying potentially meaningful information (such as speech vs nonspeech). The activation of the STS indicates a role of the ventral processing stream, responsible for building meaning from the incoming perceptual input in processing musical material incongruous with the meaning of its context.

Other studies have shown an involvement of the middle/posterior STS bilaterally in the processing of voices as opposed to other types of sounds (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000), as well as basic emotional aspects of auditory speech (Grandjean et al., 2005)[15]. It is therefore likely that the middle/posterior STS is strongly involved in extracting meaning from basic stimuli in the auditory environment. Human voices are meaningful by virtue of their specific biological significance compared to other sounds and emotional prosody is a meaningful signal used to decode the affective state of the social surroundings. A recent review suggests that the bilateral STS is a crucial structure involved in processing the social content of biologically relevant stimuli (Allison, Puce, & McCarthy, 2000). It may therefore be that under certain circumstances, that which music communicates (in the present case an emotion) is similar to the processing of meaningful speech or other biologically meaningful and relevant signals.

Thus, in answer to the question whether processing meaning in music and language is represented in a similar fashion and recruits comparable brain structures, it seems as if the two are represented in a similar but not identical fashion. Even though both types of meaning processing recruit the temporal lobe, each does so differently suggesting that the representation is not identical. Whereas processing language meaning appears to be subserved by brain regions more specifically involved in processing lexical meaning, processing meaning in music seems to activate areas predominantly related to processing communication generally and social signals in particular.

The fact that the temporal activations for both types of targets were right-lateralized and not bilateral, as one may have expected (Hickok & Poeppel, 2007) may be accounted for by several factors. The subtlety of the contrast may have obscured any bilateral activation present. In fact at lower thresholds the posterior MTG activation for word targets is also present on the left hemisphere. The fact that music was either the prime or the target may explain why there was a stronger lateralization to the right. Several studies have shown that music processing appears to be subserved by a right-dominant network of brain structures (for a review see, Zatorre, Belin, & Penhune, 2002). It is likely that this extends to processing the meaning of music. At the same time, it has been shown that processing the meaning of auditory stimuli of a more general nature (not speech specific), such as voices (Belin et al., 2000) or

[15]In this study, the STS activation was right-lateralized.

affective prosody (Grandjean et al., 2005) activates the posterior STS more strongly on the right hemisphere, which fits with the present data of the incongruous chord target activations. Thus, there is no discrepancy between the present findings and the literature on semantic processing. The strong hemispheric weighting of the present activations can be accounted for by the subtlety of the contrast, the involvement of music, as well as the functional role of the right STS, which appears to be dedicated to processing meaningful and potentially relevant acoustic information and social signals.

One additional issue to discuss is why there are differences between musically trained and untrained subjects in Experiment 5. As the present study indicates, the negativity found for the musicians is not due to a conflict at the response level, but rather at the semantic level. This suggests that musicians potentially perceive single musical stimuli as capable of communicating meaning. Musicians have to continuosly monitor the expressivity of their own musical performances as well as those of others, such as in joint music-making. Music may therefore come to have a meaning of its own, dependent on what musicians express and the intentions that it conveys. The meaning of a musical stimulus is hence directly related to the increased effort at expressing a variety of emotions through musical performance, as well as decoding this expression in the performance of others. Given the role of the STS in social perception generally, it may be that emotions expressed by music are understood as communicative signals of musical intentions when reflecting on one's own or listening to others' performances. It is therefore predicted that an fMRI study with musically untrained participants would elicit activations in the right IFG, as this signals the conflict at the basic response level, but not in the STS, as this signals the meaning derived from musical stimuli as communicative signals.

10.3 General Discussion of Experiments 5 & 6

The perceding two experiments were designed in order to test whether processing the meaning of music is represented in a similar fashion to the processing of meaning in language. An EEG study was conducted with an affective priming paradigm similar to the one used in the first three experiment reported in this dissertation, with the difference that words were used as primes and chords as targets. Having previously shown that single chords can influence the subsequent processing of affective meaning expressed in words, it was hypothesised that the meaning of single chords may therefore be represented in a similar manner to that of language meaning. The experiment was conducted using both musically trained and untrained participants and behavioural measures and ERPs were analysed.

It was found that for incongruous chord target processing only musically trained participants showed a component similar to the one reported for incongruous word target processing, namely an N400. This was accompanied by a behavioural effect as evidenced in prolonged evaluation responses to incongruous chord targets. Musically untrained participants showed neither the behavioural nor the ERP effect. As a result a second study was conducted only with musicians using fMRI to find out whether the neural correlates of the word target N400 and the chord target N400 may be comparable. Using the same paradigm as in the previous two experiments (Experiments 1 and 5) in two separate blocks it was found that whereas incongruous word targets preferentially engaged the left anterior insula and the right MTG, incongruous chord targets activated the right inferior frontal gyrus and the right posterior STS more strongly.

The findings suggest that whereas the meaning of music and of language may be superficially represented in a similar fashion, because both types of incongruous items elicit an N400, an ERP which has been taken to reflect the processing of meaning (Kutas & Hillyard, 1980), there appear to be distinct cortical structures underlying their processing. The processing of words which were affectively incongruous with the preceding chord activated the MTG (BA 21/37), a region which has been reported in previous studies on processing music meaning (Koelsch et al., 2004) and which has been specifically implicated to subserve the mapping of perceptual information to its semantic meaning (Hickok & Poeppel, 2007). Processing chords incongruous with the affect of communicated words acticated the STS, which has been reported for processing auditory stimuli meaningful on a more basic or general level (e.g. human voices, emotional prosody). These findings are compatible with various interpretations of the functions of MTG and STS and most importantly suggest that the meaning of words and language in general will presumably consistently activate areas which are involved in mapping input to lexico-semantic meaning, whereas the meaning of single chords activates areas dedicated to processing the significance of events in the auditory environment. That one may have found a *locus* of where meaning is processed in the brain is too bold a claim to make. It is by now accepted scientific canon that the respective meanings of all things potentially meaningful (e.g. words, prosody, sounds, faces) are presumably distributed widely over several cortical regions, but that some areas are recruited for mapping the perceptual input onto its acquired meaning.

The differences between musically trained and untrained participants need to be interpeted in the light of the fMRI data. It was possible to conclude on the basis of the activations found that differences were not a function of a basic response-conflict, which the musicians may have been more influenced by. Much rather, musical signals appear to have a special significance to people with musical training. Whereas this may be due to a more generally increased exposure to music with training, it may also be that the perception of music is more closely linked to the production of the signal and therefore other people responsible for its production. The STS has been shown to be functionally linked to the processing of social signals (Allison et al., 2000). As a result of the considerable training in orchestral- and ensemble playing which is a core aspect of the curriculum undertaken at conservatories, musicians may process musical signals no longer decoupled from the signaller but combined, as a result of which music may take on an increased social significance. This has to be treated as a working hypothesis, which requires further empirical study with greater care for the individual differences in time spent making music together, to develop the idea further.

Chapter 11

Summary and General Discussion

The main objective of this dissertation was to provide an empirical basis for theoretical assumptions that have been made on the nature of meaning in music. Several mechanisms by means of which meaning can arise in music have been hypothesised in the literature, for which until now there has been no evidence. Two of these mechanisms were focussed on: (1) The relationship between emotions expressed by music and meaning using several musical parameters, as well as (2) the role of tension-resolution patterns in establishing musical meaning. In addition, the neural representations of meaning in both music and language were studied, as this could inform on how similar meaning in music is to that expressed by language. These findings will now be discussed also with regards to the presence and absence of observed group differences between musically trained and untrained participants. This will be followed by some caveats and considerations impinging on the present data and end with some perspectives on future work that should be conducted in the present field.

11.1 Meaning and emotion

Possibly the most predominant aspect of musical communication lies in the expression of various emotions (Juslin & Sloboda, 2003; Krumhansl, 1997). It has been proposed that the musical expression of an emotion is understood as meaningful communication (Meyer, 1956; Swain, 1997; Koelsch et al., 2004). It has also been argued that single features of the musical input are capable of emotional expression and the communication of meaning (Koelsch & Siebel, 2005). The first three experiments presented here connect these two propositions, by varying single musical elements in the affect which they express on either a pleasant-unpleasant or happy-sad dimension, and studying their effects on subsequently presented target words either congruous or incongruous with the preceding expressed musical affect.

The first experiment varied single chords in terms of harmonic roughness, whereby chords either sounded highly consonant and therefore pleasant or highly dissonant and therefore unpleasant (Blood et al., 1999; Koelsch et al., 2006; Zentner & Kagan, 1996). These chords were used as primes after which emotional target words were presented, which either matched or did not match with the emotion expressed by the chord prime. It was hypothesised that if the emotion expressed by harmonic roughness is meaningful to listeners this should become apparent in both the speed at which target words are evaluated as well as in the recorded ERPs. It was found that words mismatched in valence were evaluated slower than matched target words. This mismatch was accompanied by an N400, which can be seen as a classic indicator of semantic processing. This study therefore provided the first link between the expression of a musical emotion and it being understood as meaningful by listeners. Musical expertise modulated neither the behavioural nor the electrophysiological responses suggesting that the recognition of an emotion expressed by music and processing this as a meaningful event occurs irrespective of training.

The second and third experiment built on the data obtained in the first experiment but using different musical parameters. Koelsch and Siebel (2005) argue that harmonic intervals in music can express emotions and convey meaning. The second experiment varied single chords in terms of harmonic intervals, which involved manipulating only the third step of the scale of the same chord to produce one major and one minor version. Because the difference entails only a semitone, processing this difference was argued to involve the fine-grained analysis of pitches and harmonic intervals. Major and minor music have been associated with the expression of happiness and sadness respectively (Hevner, 1935; Gagnon & Peretz, 2003; dalla Bella et al., 2001). The same paradigm as in the first experiment was employed, whereby chords varying in harmonic intervals (and therefore emotional expression) were used as primes, followed by emotional target words. It was hypothesised that if the emotion expressed by harmonic intervals is meaningful to listeners this should become apparent in the speed at which target words are evaluated as well as be reflected in the recorded ERPs. Similar to Experiment 1 it was found that words mismatched in valence were evaluated slower than matched target words, which in turn was accompanied by an N400 and which again was taken as evidence for semantic processing. This study therefore provided further evidence with regards to a relationship between emotional expression and meaning in music. In addition it demonstrated that a change as small as a semitone is already sufficient to convey affective meaning.

The third study extended the data obtained in Experiments 1 and 2 by investigating the emotional expression and affective meaning of instrumental timbre. This was also predicted by Koelsch and Siebel (2005) to link with emotion and meaning. Even though there has been no research on a link between instrumental timbre and emotion there is some evidence from the literature on prosody suggesting that certain basic acoustical parameters (e.g. acoustic roughness, pitch, amplitude) link to expressed vocal affect (Banse & Scherer, 1996). The third experiment varied major chords in instrumental timbre, one sounding more and the other sounding less pleasant. An analysis of each timbre showed that they differed in terms of attack time as well as the spectral centroid suggesting these features to be related to the perception of pleasantness in instrumental timbre. Using the same kind of affective paradigm as in the first two studies, it was hypothesised that if the emotion expressed by instrumental timbre is meaningful to listeners, this should become apparent in the speed at which target words are evaluated as well as be reflected in the recorded ERPs. Similar to the first two experiments it was found that words mismatched in valence were evaluated slower than matched target words, which in turn was accompanied by an N400 and which once more was interpreted as evidence for semantic processing. This study provides the first evidence that instrumental timbre can communicate emotions. In additon to this it demonstrates that instrumental timbre conveys affective meaning, which also cements the evidence from the first two studies on a relationship between emotional expression and affective meaning in music.

This set of studies supports a link between emotion and meaning in music. They represent the first piece of evidence to do so and seeing that these findings could be confirmed using several different musical parameters, it can be argued with considerable certainty that emotion appears to be a salient pathway to establishing meaning in music. This therefore confirms previous notions with regards to a relationship between meaning and emotion in music (Meyer, 1956; Sloboda, 1986; Swain, 1997; Koelsch et al., 2004; Koelsch & Siebel, 2005).

Music has been shown to be capable of a vast range of different and subtle emotions (for a comprehensive overview, see Juslin & Sloboda, 2003). Presently, a very limited subset of these was used, which are presumably perceived categorically (pleasant/unpleasant; happy/sad). This can be seen as the foundation for an increasingly refined outline of affective meaning taking more subtle musical emotions into account. We would predict that as long as these emotions are perceived and recognised, this will lead directly to an understanding of their meaning.

It is assumed that a single chord can signal a particular emotion, which when identified is categorised into a pleasant or an unpleasant event. From this process of categorisation, meaning is derived in that listeners understand that musical events belong to an affective category. Therefore emotions signalled by musical events can be said to reliably refer to a psychological construct unanimously understood by those familiar with the musical idiom. The mechanism involved in recognising the communication of an emotion is presumably a highly important evolutionary strategy. The correct identification of emotions expressed by conspecifics must have signficantly increased the chances for survival (Adolphs, 2002). Music may not normally be used to explicitly communicate an emotion in day-to-day interactions, however it is plausible that general mechanisms required for the accurate recognition of emotions are involved. Brain structures such as the amygdala and the orbitofrontal cortex have been implicated in the recognition of emotions, regardless of whether expressed visually (face) or acoustically (voice; for a review, see Adolphs, 2002). This initial recognition may then lead to further processing with regards to the emotional meaning, taking other contextual cues for disambiguation into account. For processing vocal affect, a strongly right-lateralised network of higher cortical brain areas, including frontoparietal circuits (Adolphs, Damasio, & Tranel, 2002), as well as the superior temporal sulcus (Grandjean et al., 2005) have been implicated. There is some tentative evidence that such a system may also be in place for recognising emotion portrayed by scary music (Gosselin et al., 2005; Gosselin, Peretz, Johnsen, & Adolphs, 2007). However more evidence is required on a wider range of musical emotions. Also it is important to begin devising paradigms which can reliably dissociate the perception and recognition of an emotion to the actual feeling of an emotion itself.

With regards to the meaning of the manipulated musical features, the present data suggest that this arises out of a link to meaning via emotional expression. Even though the model by Koelsch and Siebel (2005) also proposes that there is a direct link to meaning for each of these parameters, no inference on this can be drawn from the present studies. This is due to the nature of the paradigm, which was explicitly affective. Using an actual semantic priming paradigm would be able to tell if this hypothesised direct link to meaning is tenable or not.

The three experiments discussed above represent the first piece of evidence that emotions expressed in music are processed as meaningful communications. This provides relevant evidence in support of hypothesised pathways to meaning in music (Koelsch et al., 2004; Koelsch & Siebel, 2005) and is an important step in delieanating the scope and limits of musical communication.

11.2 Meaning and tension-resolution patterns

It has been suggested for a long time that musical tension-resolution patterns are meaningful to listeners familiar with Western music (Meyer, 1956; Lerdahl & Jackendoff, 1983; Sloboda, 1986). This idea has been echoed throughout many subsequent studies investigating the processing of these patterns, but no direct evidence has been reported. Meaning in this context needs careful clarification. In its initial formulation (Meyer, 1956), tension-resolution patterns were argued to be meaningful, because in a given musical context certain expectations arise with regards to subsequent musical events. These expectations reflect the dynamics of the build up and resolution of musical tension and imply musical consequences, which are of meaning to the listener. This has nothing to do with the kind of meaning that the musical expression of a specific emotion may have, but rather functions solely by virtue of referring to the musical context[16]. This has therefore been argued to reflect a kind of non-referential musical meaning (Sloboda, 1986), which does not refer to anything outside the musical context (unlike an association). Even though many studies have investigated the processing of various instances of tension-resolution

[16]The distinction between referential and non-referential (abstract patterns) meaning in music is important to note and outlined in chapter 3.2.

patterns, findings have not been able to unequivocally demonstrate the link to musical meaning.

Experiment 4 of this dissertation used a paradigm in which harmonic expectations were either violated or confirmed (thereby functioning like tension-resolution patterns). At the same time sentences were presented, which were either correct, syntactically incorrect, or semantically implausible. Assuming that similar processes in language and music recruit the same neural mechanisms (Patel, 2003), it was hypothesised that if tension-resolution patterns convey abstract meaning, the N500, a specific ERP to harmonic expectancy violations and highly reminiscient of the N400 found for processing semantic aspects of language (Koelsch et al., 2000), ought to be reduced when a semantically implausible sentence is simultaneously presented. The correct and the syntactically incorrect sentences on the other hand ought to have no significant effect on this component. It was found that the N500 was reduced significantly when semantically implausible sentences were presented and more so than when correct or syntactically incorrect sentences occurred. Especially the syntactically incorrect sentences served as an adequate control condition for general cognitive mechanisms, such as attention towards a deviant stimulus, working memory demands, as well as the size and time course of the associated ERP components. The evidence therefore strongly suggests that language and music appear to share processing resources not just dedicated to structural or syntactic processing (Patel, 2003; Koelsch et al., 2005), but also to processing meaning. In addition, this study provides a very crucial piece of evidence needed to say with certainty that musical tension-resolution patterns do indeed represent a route to meaning in music (Koelsch & Siebel, 2005; Steinbeis & Koelsch, in press).

There are two ways of accounting for the data of Experiment 4. The first entails an analogy to language processing whereby meaning can be derived from linguistic utterances as a function of syntactic rules, but in the absence of lexical information (Ullman, 2001). Chord sequences are processed according to certain rules (Koelsch et al., 2000; Patel, 2003), but there is no semantic information. Applying rules to the incoming perceptual input leads to establishing relationships between the constituent elements, which are often hierarchical (Patel, 2003), but more importantly reflect expectations that can be considered to be meaningful to the perceiver. This type of meaning is therefore unlike semantic meaning, but rather entails the kind of understanding derived from the application of learned rules to stimuli of the environment and within a familiar context. Whereas the N400 has been argued to be sensitive to contextual integration and long-term semantic memory (Kutas & Federmeier, 2000), the N500 appears to be sensitive to contextual integration only. Thus the kind of meaning tension-resolution patterns share with the one embodied by meaning in language is at the level of contextual integration rather than long-term semantic memory.

A second explanation is that the meaning inherent to tension-resolution patterns is a result of associated affective responses to violated musical expectancies. It has been shown that violating harmonic expectations leads to subjective and physiological responses indicating emotional processing (Steinbeis et al., 2005, 2006). The model by Koelsch and Siebel (2005) predicts that emotion and meaning in music are directly connected. Thus it may be that it is the affective response resulting from the violation, which conveys the meaning of the pattern. Future work may be required to dissociate these two hypotheses.

In the preceding discussion it becomes apparent that the kind of meaning argued to inhere to tension-resolution patterns is qualitatively distinct from that of emotional expression in music. Whereas the latter represents a form of referential meaning (the music referring to an emotion understood by the listener), the former represents a form of meaning which only occurs out of reference to the musical context and its structure. Whereas emotional expression is presumably something that may be perceived relatively universally, depending on the manipulated musical parameter[17], the meaning derived from tension-resolution patterns is most likely culturally specific and only occurs in those familiar with the musical idiom.

The two sets of studies (Experiments 1-3 and Experiment 4) demonstrate that meaning may be conveyed by several aspects of music, such as referring to an emotion and abstract tension-resolution patterns. This is strong evidence for multiple pathways to meaning in music, as argued previously (Swain, 1997; Koelsch et al., 2004; Koelsch & Siebel, 2005) and confirms musicological intuitions formulated over 50 years ago (Meyer, 1956).

[17]Recent evidence suggests that certain basic emotions expressed by music may be recognised by people totally unfamiliar with Western music, suggesting the existence of musical universals (Fritz et al., in preparation).

11.3 Neural representations of meaning in music and language

In addition to finding out about specific pathways to meaning in music, it was of interest to what extent meaning inherent to music and to language may be comparable. The two domains have often been compared with regards to several characteristics (e.g. syntax, phonology; Patel, 2007). As argued at an earlier stage in this dissertation, it appears as if music shares one core characteristic of language communication, which is referentiality. The first three studies presented here support this notion, albeit on a very basic level. However, it is evident that the meaning of language and music are also fundamentally different. To get one step closer to answering how comparable the two domains really are with regards to what they can communicate, looking at how they are represented at a neural level may yield important insights[18]. As has been shown, a common practice in cognitive neuroscience is to assume that specific ERP components reflect a particular cognitive function (e.g. the N400 as an indicator of semantic processing) and that these functions are localisable in particular brain areas (i.e. indicated by activation patterns in imaging studies or deficits in lesion studies). Different processes, no matter how subtle these differences may be, are presumably reflected in small differences in neural responses (e.g. scalp topographies, amplitude or latency for ERPs; small differences in activation patterns for imaging). So to be able to see whether meaning in music and language is comparable, several studies were conducted specifically investigating and comparing neural responses to meaning of both domains.

Seeing that the N400 is considered the classic marker for semantic processing, the first step in comparing the neural representations of meaning in music and language was to see if musical targets, given the appropriate context, can also elicit an N400. Previous attempts reported in the literature have failed, which may be due to the fact that the types of context used were never semantic in nature. Experiment 5 was therefore conducted with the aim to see if, in an affective context, musical targets are capable of eliciting an N400. The same paradigm as the one employed in the first three experiments was used in Experiment 5, with the difference that affective words were now the primes and chords varying in harmonic roughness were

[18] The attempts to find an N400 for musical targets is just another example of making comparisons between the nature of communication in music and language.

the targets. It was hypothesised that if the meaning of musical targets is represented in a similar fashion to that of word targets, then they also ought to elicit an N400[19].

Both musically trained and untrained subjects partook in the study. Only musicians showed an affective priming effect, where chords mismatched in valence to preceding word primes were evaluated slower than matched target chords. This mismatch was also accompanied by a negativity between 200-400 ms as well as a later positivity. Nonmusicians showed neither a behavioural effect nor a negativity. They only showed a global positivity between 500-700 ms, comparable to that found for the musicians, whereupon it was concluded that only the negativity could be functionally related to the behavioural effect in the musicians. Given that the results for the musicians could be explained both in terms of a response conflict mechanism (responding incongruously to a prime may elicit an N200, a marker of response conflict) and a semantic conflict mechanism (N400) an additional fMRI study was conducted to tease apart these two explanations, based on differing predictions of brain activations. The first part of Experiment 6 used the same experimental protocol as in Experiment 5 with only musically trained subjects and when contrasting incongruous chords against congruous chords found an activation pattern (no ACC activation, which is the supposed source of the N200, but rather activations of the right STS) that could primarily be related to a conflict at the semantic level. It was therefore concluded that the negativity found for musicians is an N400 and reflects the processing of a violation of musical meaning. This study is therefore the first successful attempt at eliciting an N400 for musical targets, given the appropriate context. The fact that this was only found for musicians will be discussed below under the heading of group differences.

The first step towards comparing the representation of meaning in music and language appears to suggest that it may be coded in a very similar manner. In response to previous comments on evidence that music can convey meaning, by using musical primes (Koelsch et al., 2004), where it was argued that presumably the type of meaning encoded in music is represented very differently to that inherent to language (Janata, 2004), it can now be said that this difference is smaller than believed. The N400 elicited by musical targets (Experiment 5) had a different distribution as

[19]That single chords varying in harmonic roughness communicate affective meaning was already established in the first study. Therefore they were considered to be suitable for addressing whether these are capable of eliciting neural representations comparable to those found for processing language meaning.

well as latency compared to the one elicited by word targets (Experiment 1). However, this may also be due to the fact that in one experiment targets were presented auditorily and in the other visually, which has been shown to affect both latency and scalp distribution of the N400 (Holcomb & Neville, 1990, this study showed an earlier N400 onset for auditory targets, as well as global scalp distribution, much like in the present study). Thus, Experiment 5 provides some compelling evidence that the meaning of music and language is represented in a highly comparable fashion.

An additional source of information on the comparability of meaning representation in music and language, can be obtained from knowing where these representations may be located. With this in mind, an fMRI study was conducted with musically trained subjects, using the same paradigms as Experiments 1 and 5. Not only did this allow looking for the specific correlates of word target integration following chord primes (which is therefore similar to the study by Koelsch et al., 2004, predicting activations of the MTG) and chord target integration following word primes (which was also required to tease apart competing explanations of the negativity found in Experiment 5), but also a direct comparison of increased integration efforts for incongruous word targets and incongruous chord targets. It was hypothesised that if the meaning of music and language are represented in a similar fashion, then both types ought to show a comparable pattern of activations.

It was found that whereas incongruous word targets contrasted with congruous word targets activated the left anterior insular cortex and the right MTG (an activation close to the coordinates reported in the source localisation by Koelsch et al., 2004), the former presumably reflecting strategic effects of word target evaluation and the latter reflecting semantic processing costs, incongruous chord targets contrasted with congruous chord targets activated the right IFG, reflecting processing response conflict and the right STS, which may also reflect increased semantic processing costs. The ensuing discussion will focus on the temporal activations found for both word and chord targets.

Processing the increased demands of both incongruous word and incongruous chord targets activated the right temporal lobe. Whereas the activation for word targets was close to the coordinates of the source localization reported by Koelsch et al. (2004), the activation for chord targets was more superior and more posterior, activating the STS. Temporal structures have been strongly implied to be involved in semantic processing (Kiehl et al., 2002; Kuperberg et al., 2000; Newman et al., 2001; Rodd et al., 2005; Rissman et al., 2003) of language and it has been argued that the MTG/STS is part of the so-called ventral processing stream of language comprehension, which hosts a lexical interface, mapping sound to meaning (Hickok & Poeppel, 2000, 2007). Thus, there is some evidence that the meaning of music and language may be processed in similar brain regions, but not identical ones. Given that words require additional lexical processing this is not surprising[20]. The chord targets activated the right posterior STS, which is also involved in processing meaning on a more basic or general level (Belin et al., 2000; Grandjean et al., 2005). It is argued therefore that the meaning which arises out of the word targets is lexical and entails the mapping of perceptual input to its lexical referent, whereas the meaning which arises out of the chord targets reflects the general relevance (e.g. biological, social) of these stimuli to those strongly familiar with them (musicians).

In sum, it appears that under certain circumstances the meaning of music and language are represented in proximal brain structures suggesting a strong similarity between the two domains, but also decided differences, which reflect the types of meaning which each stimulus can embody.

11.4 Differences between musicians and nonmusicians

The first three experiments and Experiment 5 investigated processing the relationship of meaning and emotion and the representation of musical meaning with regards to posssible group differences between musically trained and untrained subjects. As argued before, there is substantial evidence from the neuroimaging literature on functional and anatomical differences in various aspects of music processing resulting from music training. This literature is somewhat discrepant with behavioural data that have been recorded on specific musical tasks, which seem to demonstrate that apart from slight processing advantages, the cognitive mechanisms underlying music processing are not substantially different between the two groups. With regards to processing emotion, the few studies that investigated both groups found neither differences in the recognition of an emotion (Bigand et al., 2005) nor in feeling emotions in response to music (Steinbeis et al., 2005, 2006). The only study that was carried out on processing the meaning of music included nonmusicians only (Koelsch et al., 2004). Given the discrepancy between the neuroimaging

[20]This is also reflected in the activation of the right posterior MTG by incongruous word targets.

and the behavioural data on group differences and the lack of any other studies in this area, no specific hypotheses were formulated.

The first three experiments clearly demonstrate that affective meaning communicated by single chords is processed equally by both musically trained and untrained subjects. This seems to suggest that musical expertise itself has no effect on processing the relationship between meaning and emotion in music. This lack of a difference with regards to the basic perception of an emotion appears to be in line with previous studies (Bigand et al., 2005). Experiment 5 however shows some interesting differences, where only the musicians displayed a behavioural affective priming effect, as well as an N400, neither of which weres evident for the nonmusicians. The results from Experiment 6 showed that the neural correlates of processing affectively incongruous chords are located in brain structures which are known from studies on semantic processing. More importantly however, the right STS has been shown to be involved in processing meaning generally, such as the affect encoded in acoustic speech (Grandjean et al., 2005), human voices (Belin & Zatorre, 2000), as well as the biological relevance of stimuli of a social nature (for a review, see Allison et al., 2000). The superior temporal sulcus appears to encode and process meaning, which does not have to be lexical, but which signals important communicative events occurring in the environment (i.e. presence of conspecifics and their present mood). What however is so special about musical training that would make simple musical stimuli take on that kind of significance? How can single musical events come to be interpreted as a direct communication of socially important events?

Apart from the vast number of hours that musicians invest in learning their instrument, playing finger exercises and mastering difficult musical passages, making music has a strong social component[21]. Whether in orchestras, choirs, or chamber-music groups, musicians communicate with each other through their instruments, giving simple cues, shaping musical phrases and trying their best to do the composers' intentions justice. It is plausible that by virtue of this enhanced social aspect of music-making and the explicit attempts to both decode other's musical intentions as well as express oneself musically, musical signals take on a strong social meaning, which may be reflected by the activation of the STS, an area that is known for

[21] This is still important to point out in an age, where digital music has disembodied the musical message from its signaller.

processing the meaning of social signals. This is an initial working hypothesis and future studies may want to focus on correlating the extent to which musicians play in groups with their understanding music as social signals. Whereas the perception of the performer's expression by the listener is a well-studied component of music cognition, how musicians understand each other is still very poorly researched.

With regards to Experiment 4, it is only likely that musicially trained participants would display a similar reduction in the N500 with concurrently presented semantic violations as the nonmusicians. Whereas studies have shown differences resulting from musical training in the ERAN (Koelsch et al., 2002), this was not the case for the N500. Given the data from Experiments 5 and 6, where single chords appear to be processed similarly to word meaning, it is likely that also tension-resolution patterns may be more meaningful to musically trained participants. As a result one would predict a greater interaction of the N500 with the semantic material than the musically untrained participants. These two opposing hypotheses would make good material for a follow-up study to Experiment 4.

In sum, this dissertation could show that whereas musical training does not affect processing the relationship of affective meaning between a musical feature and an emotional word, it does seem to modulate the representation of affective meaning in music. The evidence used to indicate the representation of affective meaning also suggests that for musicians, the meaning of music may be coded in a similar but not identical fashion than the meaning of language. It is evident from the data that the nature of meaning underlying verbal and musical stimuli is different by virtue of the former's explicitly semantic and the latter's potentially more social nature.

11.5 Caveats, considerations and perspectives

The present dissertation investigated several issues arising out of discussions on meaning in music. It could be shown that emotions communicated by various musical parameters can convey semantically meaningful information regardless of musical expertise. In addition empirical evidence was provided for a link between tension-resolution patterns and meaning. Thus, there is clear evidence for multiple routes to meaning in music, as has been hypothesised (Koelsch et al., 2004; Koelsch & Siebel, 2005). The final part of the thesis investigated similarities in neural representations of meaning in language and music, finding evidence for considerable comparability, but only for musically trained participants. These findings were discussed with regards to music as signals of social communication. Whereas, these findings represent new and exciting data on meaning in music, they also open up several further debates relevant to the field.

Firstly, one may question the ecological validity of the present empirical approach and the subsequent drawing of general conclusions on meaning in music. What music may mean varies enormously between people and also between any given point in time. The present findings are restricted to a set of very limited and artificial circumstances designed to specifically address a set of circumscribed empirical issues. If this has any bearing on accounting for what happens when actually listening to music is still largely open. As such, they represent first steps in the right direction, but future studies should keep in mind that music listening is a contextual event, the context of which also defines what the music may or may not mean.

Another issue, which should be adressed is the cultural specificity of meaning in music. There is some evidence that certain emotions expressed in music are universally recognised (Fritz et al., in preparation). If, as was done in the present dissertation, one is to assume that the recognition of a musical emotion is sufficient to prime subsequent word processing, similar effects as reported here are expected to be found in other cultures largely unfamiliar with Western music. On the other hand, because processing tension-resolution patterns are highly culture-specific, it is unlikely that any comparable effects would be found for other cultures given the present stimulus material. However, this is an open empirical issue and subject to further investigation.

Somewhat tied in with the previous point is the issue of individual differences with regards to musical meaning. Having shown that there is some basic agreement on the meaning and emotional expression of basic musical features, it is desirable that a more comprehensive framework is drawn up, taking the likely fact of individual differences into account. Whereas the present group differences suggest an interesting effect resulting from musical training, there may well be a host of other personality variables that come into play when processing music and deciphering its meaning.

One last point is more of a rebuttal aimed towards claims on a lack of a link between meaning and music. In a recent review on the evolution of music, Fitch (2006) compares music and language using a set of thirteen criteria, including semanticity.

It is argued that language evidently has semantic properties and music does not. This however does not mean that music does not have meaning. The present thesis has partly tried to argue that music can refer to external events, such as emotions. Even though music may rarely be used in such a fashions that does not mean that it cannot do so. Discussions around the scope and limits of meaning in music should also focus on what it *can* mean and not merely on how it is used.

Appendix A

Rating Experiment

The following text describes a rating study conducted on the stimulus material used for Experiments 1 - 3, 5 and 6, including all words and all chords. Twenty-four musically trained (mean age: 24.5 years) as well as 24 untrained subjects (mean age: 23.2 years) participated in this rating study.

Ratings of words for Experiments 1, 3, 5 and 6

A pool of 150 words (pleasant and unpleasant) were rated with regards to their emotional content and their concreteness/abstractness. From this pool, 48 words (24 pleasant and 24 unpleasant) were chosen as a result of their rating scores and for which the ratings will be reported now (see also Table A.1). First the valence ratings for all words will be presented followed by the concreteness ratings for all words.

Valence

Subjects were asked to indicate how pleasant or unpleasant they found each word to be, on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant. The ratings revealed that both pleasant and unpleasant words were perceived as such by both musicians (1.69 and 4.43 respectively) and nonmusicians (1.78 and 4.48 respectively). This difference in rating was significant and $(F(1,46) = 1876.404, p < .0001)$. There were no group differences in the ratings.

Concrete/Abstract

Subjects were asked to indicate how concrete or abstract they perceived the words to be, on a scale of 1 to 5, where 1 meant abstract and 5 meant concrete. The ratings revealed that exactly half of the pleasant words and half of the unpleasant words were judged as more abstract (2.8 and 2.79 respectively) and the other half of both pleasant and unpleasant words was judged as more concrete (3.2 and 3.11 respectively). There were no group differences in the ratings. These findings suggest that the word stimuli were well balanced in the concreteness/abstractness of content.

Ratings of words for Experiment 2

A pool of 48 emotional words (24 happy and 24 sad) was created and presented to participants (also see Table A.2). First the happy/sad ratings for all words will be presented followed by the concreteness ratings for all words.

Happy/Sad

Subjects were asked to indicate how happy or sad they found each word to be, on a scale of 1 to 5, where 1 meant happy and 5 sad. The ratings revealed that both happy and sad words were perceived as such by both musicians (1.63 and 4.2 respectively) and nonmusicians (1.57 and 4.44 respectively). This difference in rating was significant and $(F(1,46) = 2104.272, p < .0001)$. There were no group differences in the ratings.

Concrete/Abstract

Subjects were asked to indicate how concrete or abstract they perceived the words to be, on a scale of 1 to 5, where 1 meant abstract and 5 meant concrete. The ratings revealed that exactly half of the happy words and half of the sad words were judged as more abstract (2.92 and 2.88 respectively) and the other half of both happy and sad words was judged as more concrete (3.08 and 3.07 respectively). There were no group differences in the ratings. These findings suggest that the word stimuli were well balanced in the concreteness/abstractness of content.

Ratings of chords used in Experiments 1, 5 and 6

A pool of 48 chords (24 consonant and 24 dissonant) was created and presented to participants. Stimuli were rated on a pleasant/unpleasant dimension.

Valence

Subjects were asked to indicate how pleasant or unpleasant they perceived each chord to be on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant. The ratings revealed that both consonant and dissonant chords were perceived as expected by both musicians (1.69 and 4.35 respectively) and nonmusicians (1.62 and 4.36 respectively). This difference in rating was significant and $(F(1,46) = 3008.072, p < .0001)$. There were no group differences in the ratings.

Ratings of chords used in Experiment 2

A pool of 48 chords (24 major and 24 minor) was created and presented to participants. Stimuli were rated on a happy/sad dimension. To see if the minor chords are necessarily perceived as less pleasant than major chords, they were also rated with regards to their pleasantness.

Valence

Subjects were asked to indicate how pleasant or unpleasant they perceived each chord to be on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant. The ratings revealed that major and minor chords were not perceived as differing with regards to their perceived pleasantness, neither for musicians (1.69 and 1.66 respectively) nor nonmusicians (1.66 and 1.69 respectively). This difference in rating was therefore not significant for either of the two groups.

Happy/Sad

Subjects were also asked to indicate how happy or sad they perceived each chord to be on a scale of 1 to 5, where 1 meant happy and 5 sad. The ratings revealed that both major and minor chords were perceived as expected by both musicians

(2.55 and 3.62 respectively) and nonmusicians (2.6 and 3.59 respectively). This difference in rating was significant and $(F(1,46) = 563,366p < .0001)$. There were no group differences in the ratings.

Ratings of chords used in Experiment 3

A pool of 48 chords (24 major chords with pleasant and 24 major chords with unpleasant timbre) was created and presented to participants. Stimuli were rated on a pleasant/unpleasant dimension.

Valence

Subjects were asked to indicate how pleasant or unpleasant they perceived each chord to be on a scale of 1 to 5, where 1 meant pleasant and 5 unpleasant. The ratings revealed that both pleasant-timbre and unpleasant chords were perceived as such by both musicians (1.7 and 4.35 respectively) and nonmusicians (1.72 and 4.36 respectively). This difference in rating was significant and $(F(1,46) = 1244.66, p < .0001)$. There were no group differences in the ratings.

Words used in Experiments 1, 3, 5 and 6	
Pleasant	*Unpleasant*
Anmut	Abscheu
Charme	Angst
Eintracht	Böses
Freund	Ekel
Frieden	Frust
Frohsinn	Furcht
Genuss	Gefahr
Grazie	Gewalt
Gutes	Gift
Heil	Gräuel
Liebe	Grauen
Lohn	Grausen
Pracht	Groll
Reiz	Hass
Ruhe	Horror
Schatz	Leiden
Schönheit	Mord
Schutz	Schaden
Segen	Schmerz
Trost	Schreck
Wohl	Übel
Wunder	Unglück
Wunsch	Wut
Ziel	Zorn

Table A.1: Table of words used in Experiment 1,3,5 and 6

Words used in Experiment 2	
Happy	*Sad*
Erfolg	Armut
Feier	Elend
Festakt	Grab
Fete	Gram
Freude	Jammer
Frohsein	Klage
Geschenk	Krankheit
Glück	Kummer
Humor	Last
Ideal	Leid
Idol	Mühsal
Jubel	Nachteil
Jux	Not
Komik	Pech
Lachen	Pein
Lust	Plage
Party	Problem
Scherz	Qual
Spass	Sorge
Spiel	Träne
Triumph	Trauer
Vorteil	Trübsal
Witz	Verlust
Wonne	Weh

Table A.2: Table of words used in Experiment 2

Appendix B

Supplementary Methods for Experiment 4

The following table displays the sentences used for Experiment 4.

Language material used in Experiment 4		
High cloze and syntactically correct	*High cloze and syntactically incorrect*	*Low cloze and syntactically correct*
Er veranstaltet das grosse Fest	Er veranstaltet den grosse Fest	Er gibt das grosse Fest
Er trinkt das kühle Bier	Er trinkt den kühlen Bier	Er sieht das kühle Bier
Sie gewinnt das schwere Spiel	Sie gewinnt den schweren Spiel	Sie kennt das schwere Spiel
Sie verstreut das feine Salz	Sie verstreut den feinen Salz	Sie benutzt das feine Salz
Sie würzt das fade Essen	Sie würzt den faden Essen	Sie schätzt das fade Essen
Sie bepflanzt das leere Beet	Sie bepflanzt den leeren Beet	Sie bearbeitet das leere Beet
Sie schickt das schwere Paket	Sie schickt den schweren Paket	Sie erkennt das schwere Paket
Er betrachtet	Er betrachtet	Er fertigt

(...continued) Language material used in Experiment 4		
High cloze and syntactically correct	*High cloze and syntactically incorrect*	*Low cloze and syntactically correct*
das bunte Bild	den bunten Bild	das bunte Bild
Er dirigiert das berühmte Orchester	Er dirigiert den berühmten Orchester	Er leitet das berühmte Orchester
Sie klöppelt das kunstvolle Deckchen	Sie klöppelt den kunstvollen Deckchen	Sie bemerkt das kunstvolle Deckchen
Sie filtert das trübe Wasser	Sie filtert den trüben Wasser	Sie reicht das trübe Wasser
Sie putzt das schmutzige Fenster	Sie putzt den schmutzigen Fenster	Sie berücksichtigt das schmutzige Fenster
Er tapeziert das geräumte Zimmer	Er tapeziert den geräumten Zimmer	Er bevorzugt das geräumte Zimmer
Er schürt das wärmende Feuer	Er schürt den wärmenden Feuer	Er erwähnt das wärmende Feuer
Sie bezieht das gemütliche Bett	Sie bezieht den gemütlichen Bett	Sie beschreibt das gemütliche Bett
Sie jätet das wuchernde Unkraut	Sie jätet den wuchernden Unkraut	Sie vergisst das wuchernde Unkraut
Sie liest das spannende Buch	Sie liest den spannenden Buch	Sie liebt das spannende Buch
Sie windelt das kleine Kind	Sie windelt den kleinen Kind	Sie kontrolliert das kleine Kind
Er zerbricht das leere Glas	Er zerbricht den leeren Glas	Er öffnet das leere Glas
Sie buchstabiert das schwierige Wort	Sie buchstabiert den schwierigen Wort	Sie erlernt das schwierige Wort
Er kapert	Er kapert	Sie erwartet

(...continued) Language material used in Experiment 4

High cloze and syntactically correct	*High cloze and syntactically incorrect*	*Low cloze and syntactically correct*
das beladene Schiff	den beladenen Schiff	das beladene Schiff
Sie bereist das bergige Land	Sie bereist das bergigen Land	Sie befährt das bergige Land
Er hobelt das trockene Holz	Er hobelt den trockenen Holz	Er beseitigt das trockene Holz
Er schlachtet das kranke Schwein	Er schlachtet den kranken Schwein	Er bewacht das kranke Schwein
Sie siebt das weiße Mehl	Sie siebt den weißen Mehl	Sie wiegt das weiße Mehl
Er fliegt das startende Flugzeug	Er fliegt den startenden Flugzeug	Er hört das startende Flugzeug
Sie näht das sommerliche Kleid	Sie näht den sommerlichen Kleid	Sie wäscht das sommerliche Kleid
Er verschrottet das langsame Auto	Er verschrottet den langsamen Auto	Er entrostet das langsame Auto
Er schärft das lange Messer	Er schärft den langen Messer	Er beachtet das lange Messer
Er mäht das grüne Gras	Er mäht den grünen Gras	Er kaut das grüne Gras
Er fließt das neue Bad	Er fließt den neuen Bad	Er ändert das neue Bad
Sie toastet das alte Brot	Sie toastet den alten Brot	Sie knetet das alte Brot
Er baut das hohe Haus	Er baut den hohen Haus	Er beantragt das hohe Haus
Er bohrt das tiefe Loch	Er bohrt den tiefen Loch	Er beendet das tiefe Loch
Sie spült das dreckige Geschirr	Sie spült den dreckigen Geschirr	Sie empfiehlt das dreckige Geschirr

(...continued) Language material used in Experiment 4		
High cloze and syntactically correct	*High cloze and syntactically incorrect*	*Low cloze and syntactically correct*
Er schleckt das schmelzende Eis	Er schleckt den schmelzenden Eis	Er verwendet das schmelzende Eis
Er singt das fröhliche Lied	Er singt den fröhlichen Lied	Er produziert das fröhliche Lied
Er schient das gebrochene Bein	Er schient den gebrochenen Bein	Er untersucht das gebrochene Bein
Er reitet das schwarze Pferd	Er reitet den schwarzen Pferd	Er pflegt das schwarze Pferd

Table B.1: Sentences used in Experiment 4

Appendix C

Supplementary ERPs for Experiments 1-3 and 5

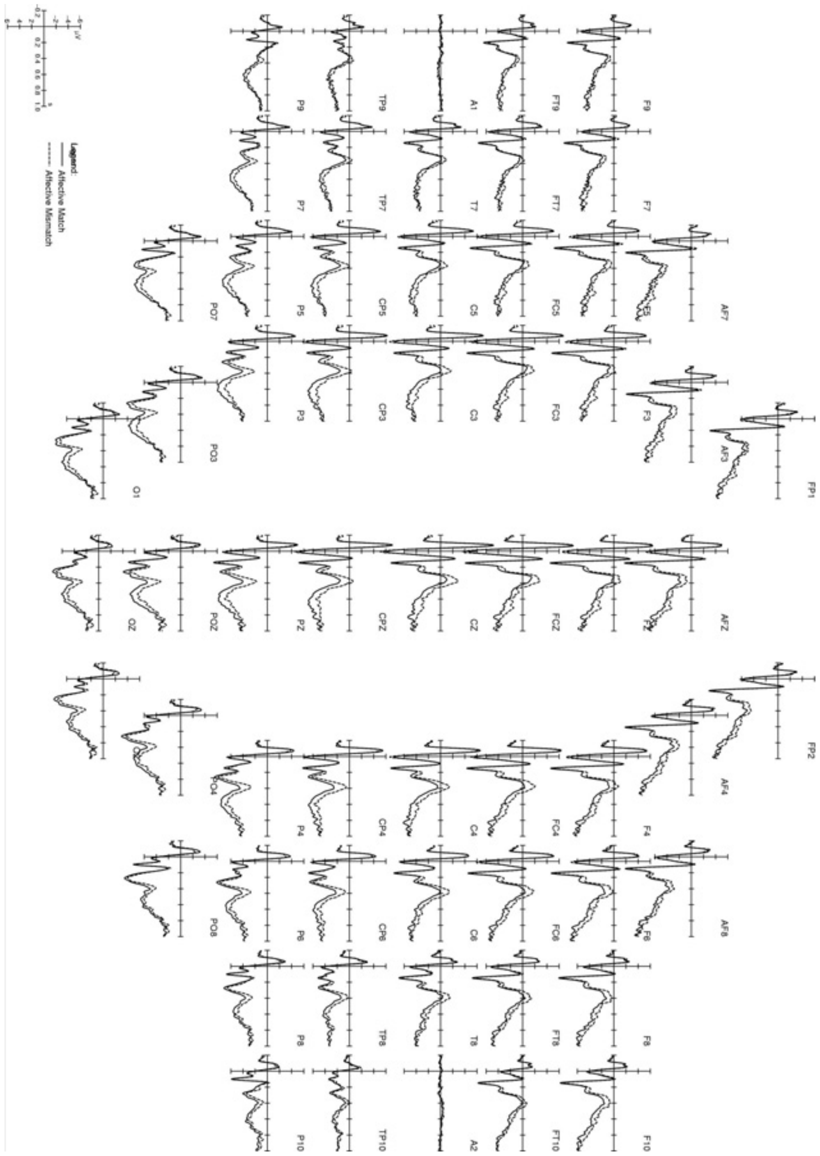

Figure C.1: Experiment 1: Supplementary ERP data for nonmusicians displayed for all electrodes.

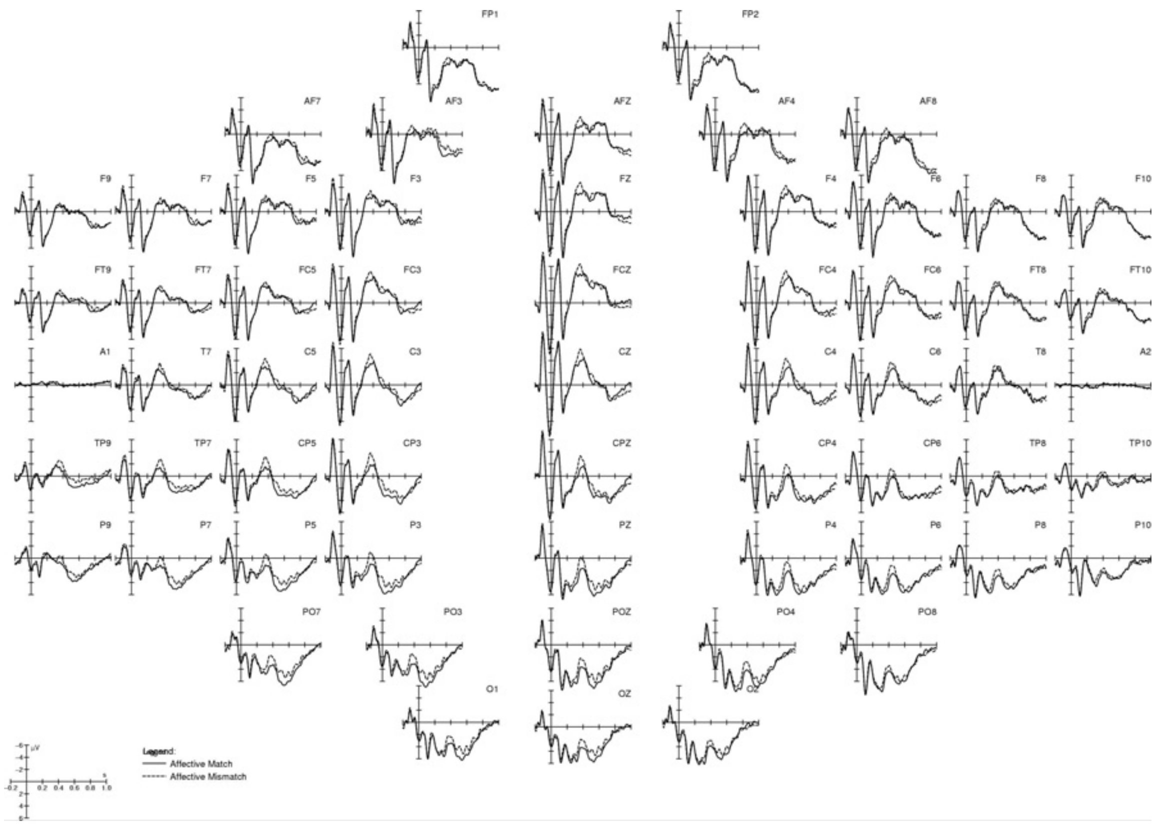

Figure C.2: Experiment 1: Supplementary ERP data for musicians displayed for all electrodes.

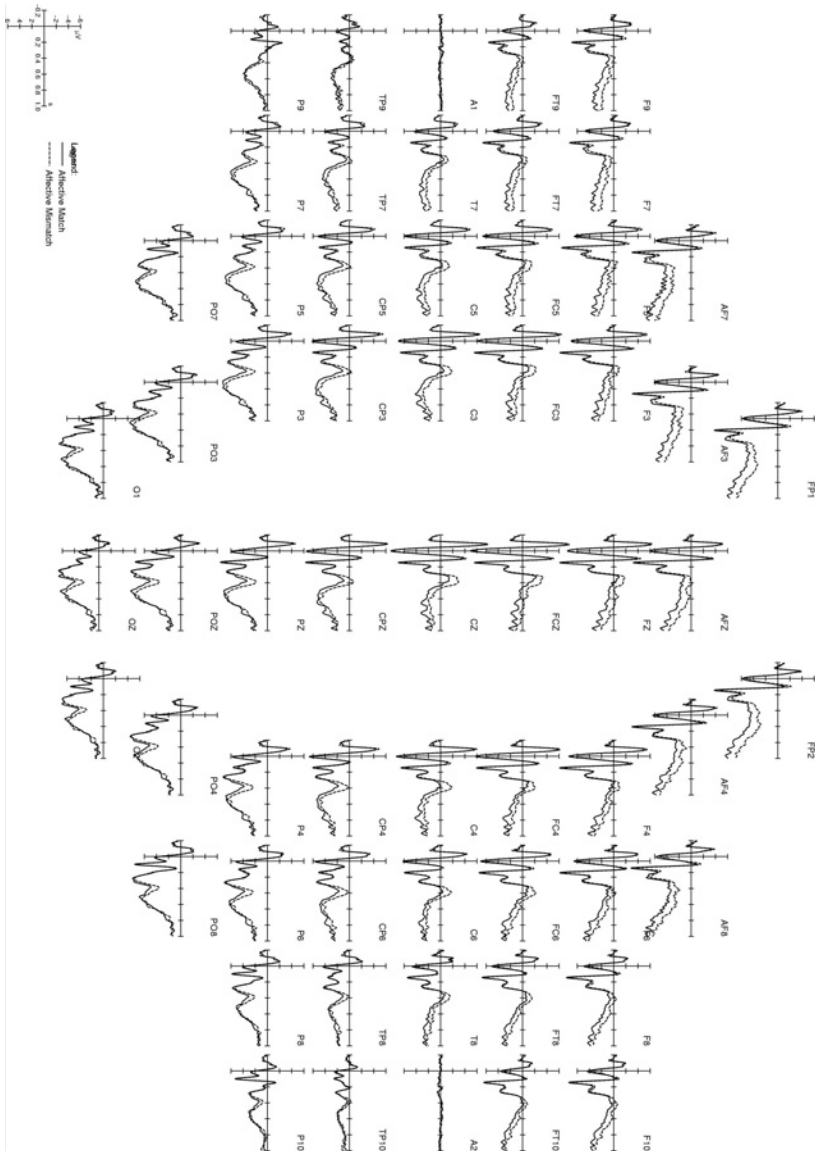

Figure C.3: Experiment 2: Supplementary ERP data for nonmusicians displayed for all electrodes.

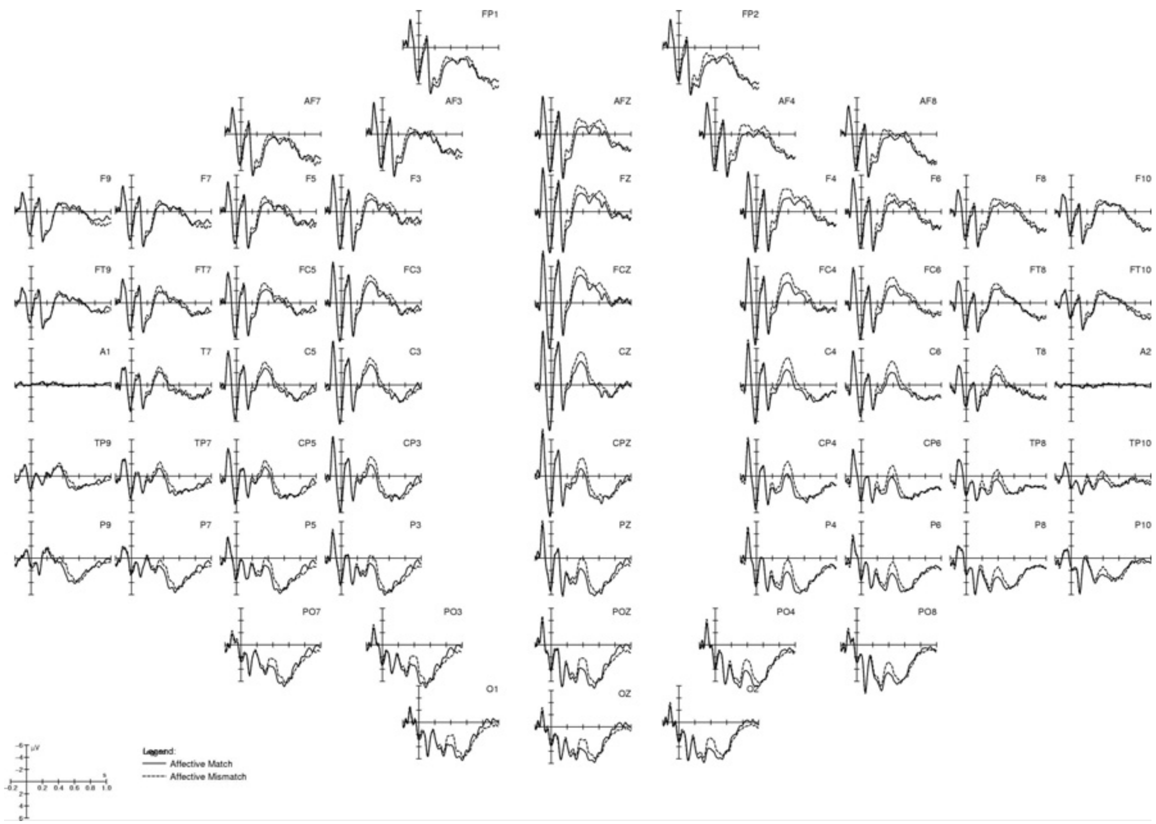

Figure C.4: Experiment 2: Supplementary ERP data for musicians displayed for all electrodes.

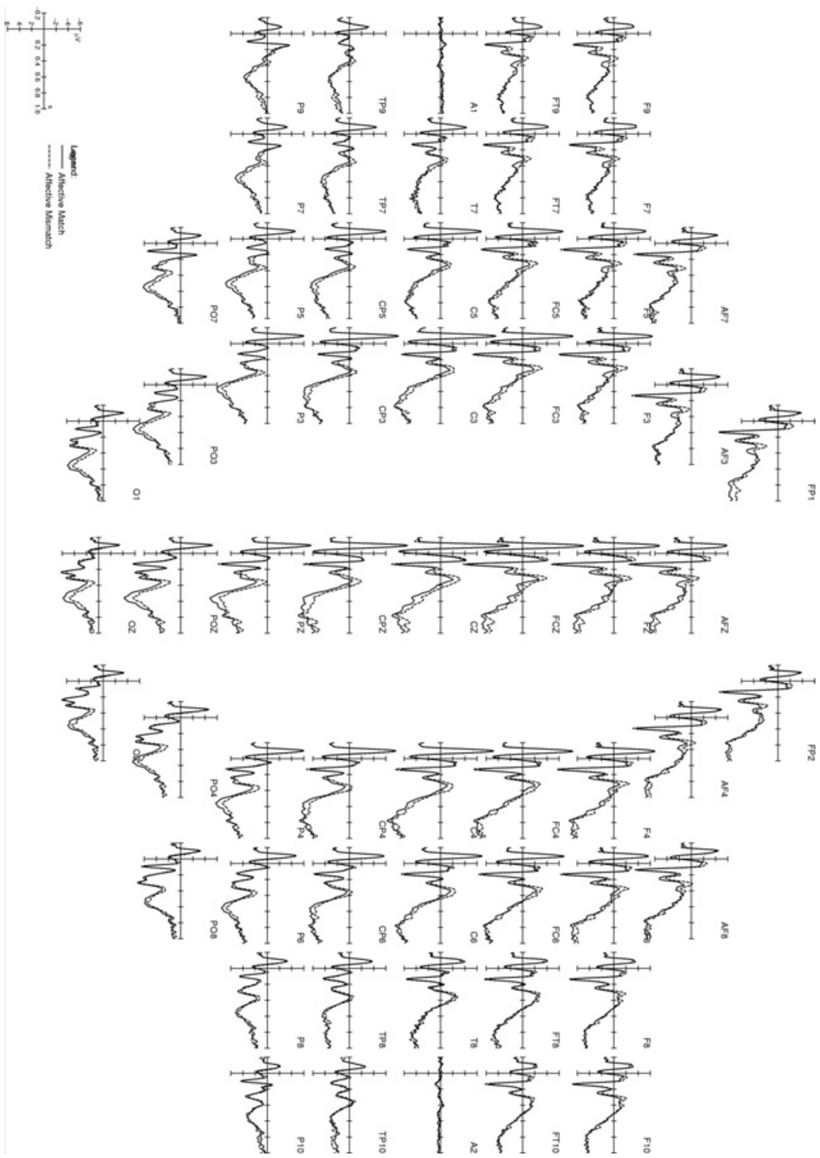

Figure C.5: Experiment 3: Supplementary ERP data for nonmusicians displayed for all electrodes.

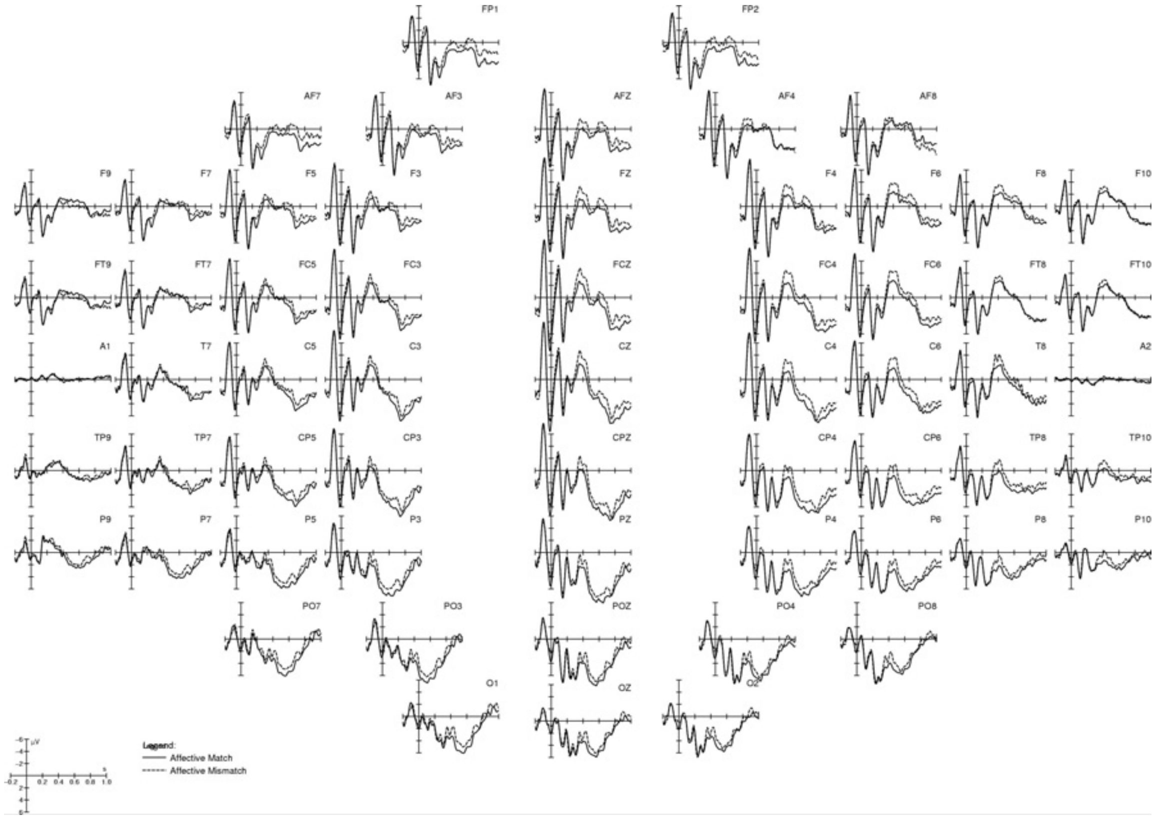

Figure C.6: Experiment 3: Supplementary ERP data for musicians displayed for all electrodes.

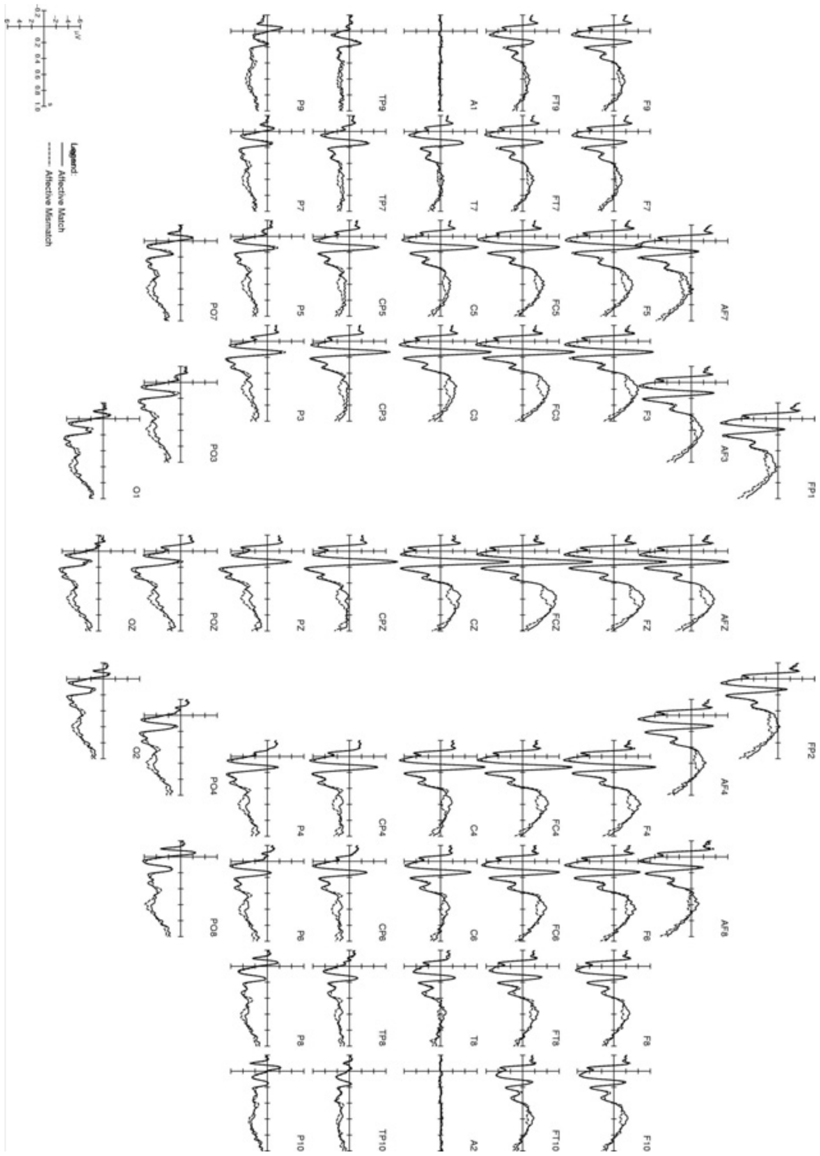

Figure C.7: Experiment 5: Supplementary ERP data for nonmusicians displayed for all electrodes.

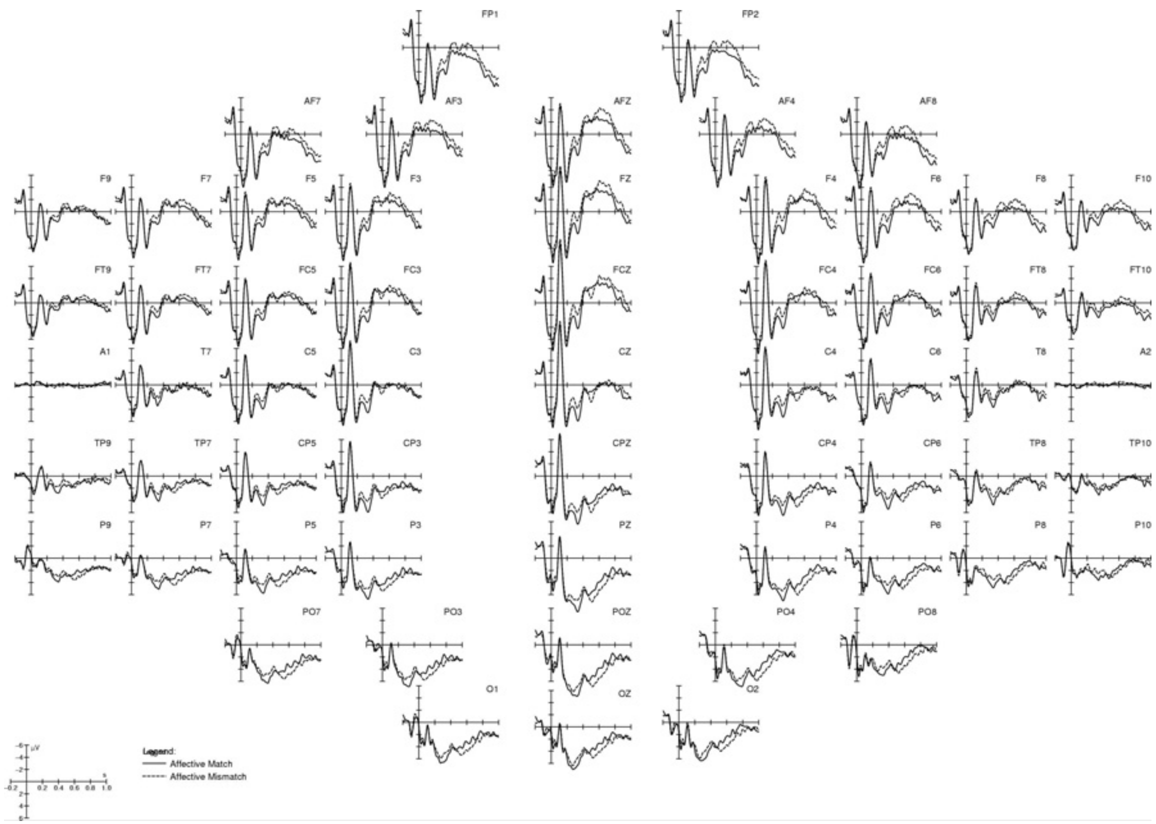

Figure C.8: Experiment 5: Supplementary ERP data for musicians displayed for all electrodes.

Bibliography

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*, 169–177.

Adolphs, R., Damasio, H., & Tranel, D. (2002). Neural systems for recognition of emotional prosody: A 3-D lesion study. *Emotion*, *2*, 23–51.

Aiello, R. (1994). Music and language: Parallels and contrasts. In R. Aiello & J. A. Sloboda (Eds.), *Musical perceptions* (p. 40-63). New York, NY: Oxford University Press.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.

Amunts, K., Schlaug, G., Jaehnke, L., Steimetz, H., Schleicher, A., Dabringhaus, A., et al. (1997). Motor cortex and hand motor skills: Structural compliance in the human brain. *Human Brain Mapping*, *5*, 206-215.

Avanzini, G., Faienza, C., Minciacchi, D., Lopez, L., & Majno, M. (Eds.). (2003). *The neurosciences and music*. Annals of the New York Academy of Sciences, Vol 999 (Special Issue).

Avanzini, G., Lopez, L., Koelsch, S., & Majno, M. (Eds.). (2005). *The neurosciences and music II: From perception to performance*. Annals of the New York Academy of Sciences, Vol. 1060 (Special Issue).

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*, 614–636.

Belin, P., & Zatorre, R. J. (2000). 'What', 'where' and 'how' in auditory cortex. *Nature Neuroscience*, *3*, 965–966.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*, 309–312.

Bernstein, L. (1976). *The unanswered question*. Cambridge, MA: Harvard University Press.

Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1278-1296.

Besson, M., & Friederici, A. (2005). Part II: Language and Music–A Comparison: Introduction. *Annals of the New York Academy of Sciences*, *1060*, 57–58.

Besson, M., & Friederici, A. D. (1998). Language and music: A comparative view. *Music Perception*, *16*, 1-9.

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, *24*, 14–25.

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences*, *930*, 232–258.

Bharucha, J., & Krumhansl, C. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, *13*, 63–102.

Bharucha, J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 403-410.

Bharucha, J., & Stoeckig, K. (1987). Priming of chords: spreading activation or overlapping frequency spectra? *Perception and Psychophysics*, *41*, 519–524.

Bigand, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music*, *18*, 45-60.

Bigand, E., Filipic, S., & Lalitte, P. (2005). The time course of emotional responses to music. *Annals of the New York Academy of Sciences*, *1060*, 429–437.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effects of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 184-197.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics*, *58*, 124–141.

Bigand, E., Perruchet, P., & Boyer, M. (1998). Implicit learning of an artificial grammar of musical timbres. *Current Psychology of Cognition*, *17*, 577-600.

Bigand, E., Poulin, B., Tillmann, B., Madurell, F., & D'Adamo, D. A. (2003). Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 159–171.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*, 100-130.

Bigand, E., Tillmann, B., Poulin-Charronnat, B., & Manderlier, D. (2005). Repetition priming: Is music special? *Quarterly Journal of Experimental Psychology A*, *58*, 1347–1375.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, *27*, 1113-1139.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*, 11818–11823.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*, 382–387.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*, 151–188.

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: From synapses to maps. *Annual Review of Neuroscience*, *21*, 149–186.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, *118*, 471–482.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, *2*, 352–363.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioural and Brain Sciences*, *22*, 77–94.

Cartmill, E. A., & Byrne, R. W. (2007). Orangutans modify their gestural signaling according to their audience's comprehension. *Current Biology*, *17*, 1345–1348.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology*, *32*, 274–285.

Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (p. 1-26). Oxford, UK: Clarendon Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428.

Cook, N. (1998). *Music: A very short introduction*. Oxford: Oxford University Press.

Cooke, D. (1959). *The language of music*. London, UK: Oxford University Press.

dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*, B1–10.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, *24*, 427–432.

de Groot, A. B. M. (1984). Primed lexical decision: Combined efefcts of the proportion of the related prime-target pairs and the stimulus-onset asychrony of prime and target. *Quarterly Journal of Experimental Psychology A*, *36*, 253-280.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*, 305–307.

Fazio, R., Sanbonmatsu, D., Powell, M., & Kardes, F. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229-238.

Fedorenko, A., Patel, A. D., Casasanto, D., Winawer, J., & Gibson, E. (2007). Structural integration in language and music: A shared system. In V. Ferreira, K. Emmorey, A. Kehler, R. Kluener, M. Polinsky, & L. Shapiro (Eds.), *20th CUNY Sentence Processing Conference, San Diego* (p. 15).

Fitch, W. T. (2005). The evolution of music in comparative perspective. *Annals of the New York Academy of Sciences*, *1060*, 29–49.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, *100*, 173–215.

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*, 377–380.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.

Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, *89*, 267–276.

Friederici, A. D., Rüschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cerebral Cortex*, *13*, 170–177.

Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *Neuroreport*, *12*, 3391–3394.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: Levels of inference and power. *Neuroimage*, *4*, 223–235.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage*, *2*, 45–53.

Fritz, T., Jentschke, S., Sammler, D., Gosselin, N., Peretz, I., & Koelsch, S. (in preparation). Universals of emotion in music.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contriubtions to "happy-sad" judgements in equition melodies. *Cognition and Emotion*, *17*, 25-40.

Gaser, C., & Schlaug, G. (2003a). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*, 9240–9245.

Gaser, C., & Schlaug, G. (2003b). Gray matter differences between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *999*, 514–517.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*, 1204–1207.

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience*, *26*, 6523–6532.

Gosselin, N., Peretz, I., Johnsen, E., & Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia*, *45*, 236–244.

Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., et al. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, *128*, 628–640.

Goydke, K. N., Altenmüller, E., Möller, J., & Münte, T. F. (2004). Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Brain Research Cognitive Brain Research*, *21*, 351–359.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*, 145–146.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*, 556–568.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, *38*, 1518–1530.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., & Marinkovic, K. (1994). Spatio-temporal stages in face and word processing I: Depth-recorded potentials in the human occipital, temporal and parietal lobes. *Journal of Physiology Paris*, *88*, 1–50.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., Marinkovic, K., Chauvel, P., et al. (1994). Spatio-temporal stages in face and word processing II: Depth-recorded potentials in the human frontal and Rolandic cortices. *Journal of Physiology Paris*, *88*, 51–80.

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex*, *10*, 69–81.

Hanslick, E. (1854/1986). *On the musically beautiful: A contribution towards the revision of the aesthetics of music* (G. Payzant, Ed.). Indianapolis: Hackett.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*, 139–151.

Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: A comparative perspective. *Nature Neuroscience*, *6*, 663–668.

Hebb, D. O. (1949). *The Organization of Behaviour: A Neurophysiological Theory*. New York: Wiley.

Hevner, K. (1935). The affective character of the major and minor modes in music. *American Journal of Psychology*, *47*, 103-118.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*, 131–138.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.

Holcomb, P., & Anderson, J. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*, *8*, 379–411.

Holcomb, P., & Neville, H. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*, 281-312.

Hund-Georgiadis, M., & von Cramon, D. Y. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Experimental Brain Research*, *125*, 417–425.

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, *7*, 153–164.

Janata, P. (2004). When music tells a story. *Nature Neuroscience*, *7*, 203–204.

Jones, M. R., & Holleran, S. (Eds.). (1991). *Cognitive bases of musical communication*. Washington D.C.: American Psychological Association.

Juslin, P. N. (2003). Communicating emotion in music performance: Review and theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and reseach* (p. 309-337). Oxford: Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (Eds.). (2003). *Music and emotion: Theory and reseach*. Oxford: Oxford University Press.

Kandel, E. R., Schwartz, J. H., & Jessel, T. M. (Eds.). (1991). *Principles of Neural Science*. Appleton and Lange: Norwalk, Conneticut.

Kelso, S. R., Ganong, A. H., & Brown, T. H. (1986). Hebbian synapses in hippocampus. *Proceedings of the National Academy of Sciences*, *83*, 5326–5330.

Kendall, R. A., & Carterette, E. C. (1990). The communication of musical expression. *Music Perception*, *8*, 129-164.

Kiehl, K. A., Laurens, K. R., & Liddle, P. F. (2002). Reading anomalous sentences: An event-related fMRI study of semantic processing. *Neuroimage*, *17*, 842–850.

Klauer, K. C., Rossnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 246–255.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*, 207–212.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Neuroimage*, *25*, 1068–1076.

Koelsch, S., Fritz, T., von Cramon, D. Y., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, *27*, 239–250.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*, 520–541.

Koelsch, S., Gunter, T., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *Neuroimage*, *17*, 956–966.

Koelsch, S., Gunter, T., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, *17*, 1565–1577.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 302–307.

Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*, 657–663.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, *10*, 1309–1313.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*, 578–584.

Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General*, *121*, 459–479.

Krumhansl, C. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension and musical ideas. *Music Perception*, *13*, 401-432.

Krumhansl, C. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psycholoy*, *51*, 336–353.

Krumhansl, C. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, *2*, 45-50.

Krumhansl, C., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*, 334-368.

Krumhansl, C., & Shepard, R. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 579–594.

Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., et al. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: An fMRI study. *Journal of Cognitive Neuroscience*, *12*, 321–341.

Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470.

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.

Kutas, M., & Hilyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.

Lamprecht, R., & LeDoux, J. (2004). Structural plasticity and memory. *Nature Reviews Neuroscience*, *5*, 45–54.

Lee, D. J., Chen, Y., & Schlaug, G. (2003). Corpus callosum: Musician and gender effects. *Neuroreport*, *14*, 205–209.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levitin, D. J., Menon, V., Schmitt, J. E., Eliez, S., White, C. D., Glover, G. H., et al. (2003). Neural correlates of auditory perception in Williams syndrome: An fMRI study. *Neuroimage*, *18*, 74–82.

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex*, *15*, 1621–1631.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*, 150–157.

Lohmann, G., Müller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., et al. (2001). Lipsia: A new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics*, *25*, 449–457.

Luria, A. R., Tsvetkova, L. S., & Futer, D. S. (1965). Aphasia in a composer (V. G. Shebalin). *Journal of Neurological Science*, *2*, 288–292.

Maess, B., Herrmann, C. S., Hahne, A., Nakamura, A., & Friederici, A. D. (2006). Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Research Cognitive Brain Research*, *1096*, 163–172.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, *4*, 540–545.

Marin, O. S. M., & Perry, D. W. (1999). Neurological aspects of music perception and performance. In D. Deutsch (Ed.), *The psychology of music* (p. 652-724). San Diego, CA: Academic Press.

Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of language processes*. Hillsdale, NJ: Lawrence Erlbaum.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*, 71–102.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal strutcure of spoken language understanding. *Cognition*, *8*, 1-71.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound: The Cognitive Psychology of Human Audition* (p. 146-198). Oxford: Oxford University Press.

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

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McClelland, J. L., & Rumelhart, D. E. (1981). An inteactive activation model of context effects in letter perception Part 1: An account of basic findings. *Psychological Review*, *88*, 375-407.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–197.

McDermott, J., & Hauser, M. D. (2005). Probing the evolutionary origins of music perception. *Annals of the New York Academy of Sciences*, *1060*, 6–16.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.

Menon, V., Levitin, D. J., Smith, B. K., Lembke, A., Krasnow, B. D., Glazer, D., et al. (2002). Neural correlates of timbre change in harmonic sounds. *Neuroimage*, *17*, 1742–1754.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, P., Mecklinger, A., Grunwald, T., Fell, J., Elger, C. E., & Friederici, A. D. (2005). Language processing within the human medial temporal lobe. *Hippocampus*, *15*, 451–459.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain Research Cognitive Brain Research*, *12*, 467–473.

Miranda, R. A., & Ullman, M. T. (in press). Double dissociation between rules and memory in music: An event-related potential study. *Neuroimage*.

Mummery, C. J., Shallice, T., & Price, C. J. (1999). Dual-process model in semantic priming: A functional imaging perspective. *Neuroimage*, *9*, 516–525.

Musch, J., & Klauer, K. C. (Eds.). (2003). *The psychology of evaluation*. Mahwah, NJ: Lawrence Erlbaum.

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*, 473–478.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226-254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (p. 264-336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Neumann, J., & Lohmann, G. (2003). Bayesian second-level analysis of functional magnetic resonance images. *Neuroimage*, *20*, 1346–1355.

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguistic Research*, *30*, 339–364.

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., et al. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience*, *12*, 120–133.

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, *372*, 260–263.

Nobre, A. C., & McCarthy, G. (1995). Language-related field potentials in the anterior-medial temporal lobe II: Effects of word type and semantic priming. *Journal of Neuroscience*, *15*, 1090–1098.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends in Neurosciences*, *24*, 283–288.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 786–803.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 331–346.

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115–138.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, *12*, 169–174.

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

Patel, A. D. (2007). Language, music, and the brain: A resource-sharing framework. In *Conference on Language and Music as Cognitive Systems, Cambridge*.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*, 717–733.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Quarterly Journal of Experimental Psychology A*, *46*, 301–325.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., et al. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, *117*, 1283–1301.

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, *389*, 495–498.

Pinker, S. (1994). *The language instinct*. New York: Nortin.

Plomp, R., & Levelt, W. J. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, *38*, 548–560.

Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: An event-related potential study. *Journal of Cognitive Neuroscience*, *18*, 1545–1554.

Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, *94*, B67–B78.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences*, *4*, 512-526.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioural Science*, *12*, 410–430.

Raffman, D. (1993). *Language, music, and mind*. Cambridge, MA: MIT Press.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408.

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*, 241–255.

Repp, B. H. (1998). A microcosm of musical expression I: Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America*, *104*, 1085–1100.

Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience*, *15*, 1160–1175.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, *15*, 1261–1269.

Rossell, S. L., Price, C. J., & Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia*, *41*, 550–564.

Rugg, M., & Coles, M. (1995). *Electrophysiology of mind: Event-related potentials and cognition*. Oxford: Oxford University Press.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception Part 2: The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*, 60–94.

Rüsseler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Münte, T. F. (2001). Event-related brain potentials to sound omissions differ in musicians and nonmusicians. *Neuroscience Letters*, *308*, 33–36.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Senghas, A., & Trueswell, J. C. (2001). The acquisition of language by children. *Proceedings of the National Academy of Sciences*, *98*, 12874–12875.

Scherer, K. R. (1989). Vocal correlates of emotional arousal and affective disturbance. In H. Wagner & A. Manstead (Eds.), *Handbook of Social Psychophysiology* (chap. 165-197). London, UK: Wiley.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*, 76-92.

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a Sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience*, *15*, 1135–1148.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*, 24–30.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Research Cognitive Brain Research*, *14*, 228–233.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Brain Research Cognitive Brain Research*, *24*, 442–452.

Schirmer, A., Zysset, S., Kotz, S. A., & Cramon, D. Y. von. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *Neuroimage*, *21*, 1114–1123.

Schlaug, G. (2001). The brain of musicians: A model for functional and structural adaptation. *Annals of the New York Academy of Sciences*, *930*, 281–299.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*, 1047–1055.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*, 688–694.

Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., et al. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience*, *8*, 1241–1247.

Schubert, E., Wolfe, J., & Tarnopolsky, A. (2004). Spectral centroid and timbre in complex, multiple instrumental textures. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the International Conference on Music Perception and Cognition, North Western University, Illinois* (p. 112-116).

Seyfarth, R. M., & Cheney, D. L. (2003a). Meaning and emotion in animal vocalizations. *Annals of the New York Academy of Sciences*, *1000*, 32–55.

Seyfarth, R. M., & Cheney, D. L. (2003b). Signallers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145-173.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, *210*, 801–803.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-190.

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2007). Language, music, and modularity: Self-paced reading time evidence for shared processing of linguistic and musical syntax. In V. Ferreira, K. Emmorey, A. Kehler, R. Kluener, M. Polinsky, & L. Shapiro (Eds.), *20th CUNY Sentence Processing Conference, San Diego* (p. 22).

Sloboda, J. A. (1986). *The musical mind*. Oxford: Oxford University Press.

Sloboda, J. A., & Lehmann, A. C. (2001). Tracking performance correlates of changes in perceived intensity of emotion during different interpretations of a Chopin Piano Prelude. *Music Perception*, *19*, 87-120.

Sollberger, B., Reber, R., & Eckstein, D. (2003). Musical chords as affective priming context in a word-evaluation task. *Music Perception*, *20*, 263-282.

Steinbeis, N., & Koelsch, S. (in press). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2005). Emotional processing of harmonic expectancy violations. *Annals of the New York Academy of Sciences*, *1060*, 457–461.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*, 1380–1393.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196.

Stravinsky, I. (1936). *An autobiography*. New York: Simon and Schuster.

Swain, J. (1997). *Musical languages*. Norton, New York.

Talairach, J., & Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - an Approach to Cerebral Imaging*. New York: Thieme Medical Publishers.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *94*, 14792–14797.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*, 885–913.

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Brain Research Cognitive Brain Research*, *16*, 145–161.

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., et al. (2006). Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *Neuroimage*, *31*, 1771–1782.

Tomasello, Call, & Hare. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*, 1063–1069.

Tulving, E., & Donaldson, W. (Eds.). (1972). *The organisation of memory*. New York: Academic Press.

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, *2*, 717–726.

van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*, 657–671.

van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes*, *8*, 485-531.

van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 394–417.

van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related potentials. *Memory and Cognition*, *18*, 380-393.

van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition*, *19*, 95–112.

van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*, 593–602.

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., & Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage*, *14*, 1302–1308.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*, 329–338.

Wallin, N. L., Merker, B., & Brown, S. (2000). *The origins of music*. Cambridge, Mass: MIT Press.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 456–469.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited–again. *Neuroimage*, *2*, 173–181.

Zatorre, R., Belin, P., & Penhune, V. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*, 37–46.

Zatorre, R., & Peretz, I. (Eds.). (2001). *The biological foundations of music*. Annals of the New York Academy of Sciences, Vol. 930 (Special Issue).

Zentner, M. R., & Kagan, J. (1996). Perception of music by infants. *Nature*, *383*, 29.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25–64.

Zwitserlood, P. (1994). Access to phonological form representations in language comprehension and production. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (p. 20-35). Hillsdale, NJ: Lawrence Erlbaum.

Zwitserlood, P. (1998). Spoken words in sentence contexts. In A. D. Friederici (Ed.), *Language comprehension: A biological perspective* (p. 71-99). Berlin-Heidelberg: Springer.

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of Variance
BA	Brodmann Area
BOLD	Blood oxygenated level dependent
cACC	caudal Anterior cingulate cortex
CC	Corpus callosum
EEG	Electroencephalogram
EKP	Ereignis-korreliertes Potential
EPI	Echo planar imaging
EPSP	Excitatory post-synaptic potential
ERAN	Early right anterior negativity
ERP	Event-related potential
FFT	Fast Fourier Transformation
FID	Free Induction Decay
fMRI	functional Magnetic Resonance Imaging
fMRT	funktionelle Magnet Resonanz Tomographie
GLM	General linear model
Hz	Hertz
IFG	Inferior frontal gyrus

LAN	Left anterior negativity
LTP	Long-term potentiation
MDS	Multidimensional scaling
MEG	Magnetoencephalogram
MMN	Mismatch negativity
MTG	Middle temporal gyrus
NMR	Nuclear Magnetic Resonance
PET	Positron Emission Tomography
RF	Radio frequency
RMS	Root Mean Square
ROI	Region of Interest
SD	Standard Deviation
SE	Standard Error
SOA	Stimulus Onset Asychrony
SSIRH	Shared syntactic integration resource hypothesis
STS	Superior temporal sulcus
TE	time echo
TR	Recylcle time

Curriculum Vitae

Name	Nikolaus Steinbeis
Date of Birth	16.04.1979
Place of Birth	München

March 2005 – present	*PhD. student*, Max-Planck Institute for Human Cognitive and Brain Sciences Leipzig
September 2003 – August 2004	*MSc* in Music Psychology, University of Keele, UK
September 1998 – August 2001	*BSc* in Psychology, University College London, UK
1995 – 1997	*International Baccalaureat*, Sevenoaks School, UK
1990 – 1995	European School, Culham, UK
	Ludwigsgymnasium, München
	Landesgymnasium Miesbach, Miesbach

Selbständigkeitserklärung

Hiermit erkläre ich, dass die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde und dass die aus fremden Quellen direkt oder indirekt übernommenen Gedanken in der Arbeit als solche kenntlich gemacht worden sind.

Nikolaus Steinbeis
Leipzig, 26. Oktober 2007

Bibliographic details

Steinbeis, Nikolaus

Investigating the meaning of music using EEG and fMRI

Universität Leipzig

209 pages, 235 references, 35 figures, 11 tables

Paper The aim of the dissertation was to investigate the nature of meaning which can be communicated by music. Several musical routes to meaning, which have been hypothesised in the literature but so far not researched, were considered and investigated, specifically the roles of emotion and tension-resolution patterns. In addition the dissertation investigated the comparability of how meaning is represented in both music and language, as indicated by associated neural signatures. The role of musical training was also investigated with regards to the emotional route as well as the representation of meaning.

The expression of emotions in music has long been considered to communicate meaningful information to listeners. This was studied in three separate experiments each using one of three musical features linked with emotional expression (harmonic roughness, harmonic intervals, instrumental timbre) and manipulating chords according to each of these. Using an affective priming paradigm, chord primes were paired with word targets, which were either congruous or incongruous in valence. Behavioural and electrophysiological responses (specifically event-related potentials, ERPs) revealed that the emotion expressed by a single chord can modulate processing the meaning of subsequently presented words, which in turn could be shown for each of the three features under investigation. No differences as a function of musical training were found.

In a further experiment the ability of tension-resolution patterns in communicating meaning was studied using an interactive paradigm with chord sequences and sentences containing different types of violations. Analysis of electrophysiological responses showed that an ERP in response to the violation of harmonic expectations could be modulated specifically by semantic anomalies occuring in simultaneously presentated sentences. This was taken to suggest that language and music share

neural resources for processing meaning and that tension-resolution patterns can communicate meaning to listeners familiar with Western music.

The comparability of how meaning is represented in both music and language was studied using an affective priming paradigm similar to Experiments 1-3 but using chords as targets and words as primes. The analysis of ERPs showed that only musically trained participants displayed an N400 in response to incongruous chord targets, comparable to that found for incongruous target words. A subsequent fMRI study revealed that whereas incongruous word targets activate the posterior middle temporal gyrus (MTG) involved in mapping perceptual input to its lexico-semantic meaning, incongruous chords activated the posterior superior temporal sulcus (STS) involved in identifying potentially meaningful information in one's surroundings. The data speak for a certain similarity in the way meaning is represented in language and music, but also highlight important functional differences.

Referat Das Ziel der vorliegenden Dissertation war die Untersuchung bisheriger Annahmen zur Bedeutung von Musik. In der Literatur bestehen seit längerer Zeit Hypothesen zu den möglichen Mechanismen durch welche Musik Bedeutung übermittlen kann, die bis dato noch nicht überprüft worden sind. Die Dissertation konzentriert sich vor allem auf zwei dieser Mechanismen, (1) den musikalischen Ausdruck von Emotionen oder Stimmungen, sowie (2) den Auf- und Abbau von musikalischen Spannungsbögen. Zusätzlich wurde anhand neuronaler Korrelate untersucht, ob und wie sich die Repräsentationen der Bedeutung von Musik und Sprache miteinander vergleichen lassen. Einzelne Fragestellungen wurden ausserdem im Hinblick auf den Einfluss musikalischen Trainings untersucht.

Ob der musikalische Ausdruck von Emotionen auch Bedeutung kommunizieren kann wurde anhand von 3 Experimenten überprüft. In diesen Studien wurden einzelne musikalische Parameter (Experiment 1: harmonische Rauhigkeit, Experiment 2: harmonische Intervalle, Experiment 3: instrumentale Klangfarbe) von Akkorden manipuliert um Unterschiede im emotionalen Ausdruck (angenehm oder unangenhem) zu erzielen. Anschließend wurden die Akkorde mit Wörtern gepaart, die mit der Emotion entweder kongruent waren oder nicht. Sowohl die Verhaltensdaten als auch die Analyse der Ereignis-korrelierten Potentiale (EKPs) der drei Experimente zeigten, dass der emotionale Ausdruck einzelner Akkorde die Bedeutungsverarbeitung darauffolgender Wörter beeinflussen konnte. Diese Effekte waren unab-

hängig vom musikalischen Training und sprechen dafür, dass der musikalische Ausdruck von Emotion Bedeutung kommunizieren kann.

Eine weitere Studie untersuchte ob der Auf- und Abbau von musikalischen Spannungsbögen ebenfalls Bedeutung vermittlen kann. Unter Verwendung eines interaktiven Paradigmas, in dem musikalische Sequenzen und Sätze gleichzeitig präsentiert wurden, konnte gezeigt werden, dass sich ein musikbezogenes EKP spezifisch durch die semantische Verletzung eines gleichzeitig präsentierten Satzes modulieren ließ. Aus diesem Ergebnis ließ sich schließen, dass sich Musik und Sprache neuronale Ressourcen für die Verarbeitung von Bedeutung teilen, sowie dass der Auf- und Abbau von musikalischen Spannungsbögen von musikalischer Bedeutung ist.

In zwei weiteren Studien wurden die Repräsentationen der Bedeutung von Musik und Sprache im Hinblick auf deren Vergleichbarkeit untersucht. Dies erfolgte anhand einer modifizierten Version des Paradigmas der ersten drei Studien, wobei nicht mehr Wörter die Zielreize waren sondern Akkorde. Nur die Verhaltensdaten und EKPs der Gruppe mit musikalischem Training zeigte einen Effekt, der mit dem auf Wort-Zielreize gefundenen Effekt der ersten drei Studien vergleichbar war. Eine fMRT-Studie konnte zeigen, dass inkongruente Wort-Zielreize als auch inkongruente Akkord-Zielreize den rechten Temporallappen aktivierten, davon jedoch unterschiedliche Regionen. Diese Ergebnisse implizieren dass oberflächliche Ähnlichkeiten zwischen den Repräsentationen musikalischer Bedeutung und der Bedeutung von Sprache bestehen, jedoch auch signifikante Unterschiede vorhanden sind.

MPI Series in Human Cognitive and Brain Sciences:

1 Anja Hahne
Charakteristika syntaktischer und semantischer Prozesse bei der auditiv Sprachverarbeitung: Evidenz aus ereigniskorrelierten Potentialstudien

2 Ricarda Schubotz
Erinnern kurzer Zeitdauern: Behaviorale und neurophysiologische Korrelate einer Arbeitsgedächtnisfunktion

3 Volker Bosch
Das Halten von Information im Arbeitsgedächtnis: Dissoziationen langsamer corticaler Potentiale

4 Jorge Jovicich
An investigation of the use of Gradient- and Spin-Echo (GRASE) imaging for functional MRI of the human brain

5 Rosemary C. Dymond
Spatial Specificity and Temporal Accuracy in Functional Magnetic Resonance Investigations

6 Stefan Zysset
Eine experimentalpsychologische Studie zu Gedächtnisabrufprozessen unter Verwendung der funktionellen Magnetresonanztomographie

7 Ulrich Hartmann
Ein mechanisches Finite-Elemente-Modell des menschlichen Kopfes

8 Bertram Opitz
Funktionelle Neuroanatomie der Verarbeitung einfacher und komplexer akustischer Reize: Integration haemodynamischer und elektrophysiologischer Maße

9 Gisela Müller-Plath
Formale Modellierung visueller Suchstrategien mit Anwendungen bei der Lokalisation von Hirnfunktionen und in der Diagnostik von Aufmerksamkeitsstörungen

10 Thomas Jacobsen
Characteristics of processing morphological structural and inherent case in language comprehension

11 Stefan Kölsch
Brain and Music
A contribution to the investigation of central auditory processing with a new electrophysiological approach

12 Stefan Frisch
Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen

13 Markus Ullsperger
The role of retrieval inhibition in directed forgetting – an event-related brain potential analysis

14 Martin Koch
Measurement of the Self-Diffusion Tensor of Water in the Human Brain

15 Axel Hutt
Methoden zur Untersuchung der Dynamik raumzeitlicher Signale

32 André J. Szameitat
Die Funktionalität des lateral-präfrontalen Cortex für die Verarbeitung von Doppelaufgaben

33 Susanne Wagner
Verbales Arbeitsgedächtnis und die Verarbeitung ambiger Wörter in Wort- und Satzkontexten

34 Sophie Manthey
Hirn und Handlung: Untersuchung der Handlungsrepräsentation im ventralen prämotorischen Cortex mit Hilfe der funktionellen Magnet-Resonanz-Tomographie

35 Stefan Heim
Towards a Common Neural Network Model of Language Production and Comprehension: fMRI Evidence for the Processing of Phonological and Syntactic Information in Single Words

36 Claudia Friedrich
Prosody and spoken word recognition: Behavioral and ERP correlates

37 Ulrike Lex
Sprachlateralisierung bei Rechts- und Linkshändern mit funktioneller Magnetresonanztomographie

38 Thomas Arnold
Computergestützte Befundung klinischer Elektroenzephalogramme

39 Carsten H. Wolters
Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain

40 Ansgar Hantsch
Fisch oder Karpfen? Lexikale Aktivierung von Benennungsalternative bei der Objektbenennung

41 Peggy Bungert
Zentralnervöse Verarbeitung akustischer Informationen
Signalidentifikation, Signallateralisation und zeitgebundene Informationsverarbeitung bei Patienten mit erworbenen Hirnschädigungen

42 Daniel Senkowski
Neuronal correlates of selective attention: An investigation of electro-physiological brain responses in the EEG and MEG

43 Gert Wollny
Analysis of Changes in Temporal Series of Medical Images

44 Angelika Wolf
Sprachverstehen mit Cochlea-Implantat: EKP-Studien mit postlingual ertaubten erwachsenen CI-Trägern

45 Kirsten G. Volz
Brain correlates of uncertain decisions: Types and degrees of uncertainty

46 Hagen Huttner
Magnetresonanztomographische Untersuchungen über die anatomische Variabilität des Frontallappens des menschlichen Großhirns

47 Dirk Köster
Morphology and Spoken Word Comprehension: Electrophysiological Investigations of Internal Compound Structure

48 Claudia A. Hruska
Einflüsse kontextueller und prosodischer Informationen in der auditorischen Satzverarbeitung: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

49 Hannes Ruge
Eine Analyse des raum-zeitlichen Musters neuronaler Aktivierung im Aufgabenwechselparadigma zur Untersuchung handlungssteuernder Prozesse

50 Ricarda I. Schubotz
Human premotor cortex: Beyond motor performance

51 Clemens von Zerssen
Bewusstes Erinnern und falsches Wiedererkennen:
Eine funktionelle MRT Studie neuroanatomischer Gedächtniskorrelate

52 Christiane Weber
Rhythm is gonna get you.
Electrophysiological markers of rhythmic processing in infants with and without risk for Specific Language Impairment (SLI)

53 Marc Schönwiesner
Functional Mapping of Basic Acoustic Parameters in the Human Central Auditory System

54 Katja Fiehler
Temporospatial characteristics of error correction

55 Britta Stolterfoht
Processing Word Order Variations and Ellipses: The Interplay of Syntax and Information Structure during Sentence Comprehension

56 Claudia Danielmeier
Neuronale Grundlagen der Interferenz zwischen Handlung und visueller Wahrnehmung

57 Margret Hund-Georgiadis
Die Organisation von Sprache und ihre Reorganisation bei ausgewählten, neurologischen Erkrankungen gemessen mit funktioneller Magnetresonanztomographie – Einflüsse von Händigkeit, Läsion, Performanz und Perfusion

58 Jutta L. Mueller
Mechanisms of auditory sentence comprehension in first and second language: An electrophysiological miniature grammar study

59 Franziska Biedermann
Auditorische Diskriminationsleistungen nach unilateralen Läsionen im Di- und Telenzephalon

60 Shirley-Ann Rüschemeyer
The Processing of Lexical Semantic and Syntactic Information in Spoken Sentences:
Neuroimaging and Behavioral Studies of Native and Non-Native Speakers

61 Kerstin Leuckefeld
The Development of Argument Processing Mechanisms in German.
An Electrophysiological Investigation with School-Aged Children and Adults

62 Axel Christian Kühn
Bestimmung der Lateralisierung von Sprachprozessen unter besondere Berücksichtigung des temporalen Cortex, gemessen mit fMRT

63 Ann Pannekamp
Prosodische Informationsverarbeitung bei normalsprachlichem und deviantem Satzmaterial: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

64 Jan Derrfuß
Functional specialization in the lateral frontal cortex: The role of the inferior frontal junction in cognitive control

65 Andrea Mona Philipp
The cognitive representation of tasks
Exploring the role of response modalities using the task-switching paradigm

66 Ulrike Toepel
Contrastive Topic and Focus Information in Discourse – Prosodic Realisation and Electrophysiological Brain Correlates

67 Karsten Müller
Die Anwendung von Spektral- und Waveletanalyse zur Untersuchung der Dynamik von BOLD-Zeitreihen verschiedener Hirnareale

68 Sonja A.Kotz
The role of the basal ganglia in auditory language processing: Evidence from ERP lesion studies and functional neuroimaging

69 Sonja Rossi
The role of proficiency in syntactic second language processing: Evidence from event-related brain potentials in German and Italian

70 Birte U. Forstmann
Behavioral and neural correlates of endogenous control processes in task switching

71 Silke Paulmann
Electrophysiological Evidence on the Processing of Emotional Prosody: Insights from Healthy and Patient Populations

72 Matthias L. Schroeter
Enlightening the Brain – Optical Imaging in Cognitive Neuroscience

73 Julia Reinholz
Interhemispheric interaction in object- and word-related visual areas

74 Evelyn C. Ferstl
The Functional Neuroanatomy of Text Comprehension

75 Miriam Gade
Aufgabeninhibition als Mechanismus der Konfliktreduktion zwischen Aufgabenrepräsentationen

76 Juliane Hofmann
Phonological, Morphological, and Semantic Aspects of Grammatical Gender Processing in German

77 Petra Augurzky
Attaching Relative Clauses in German – The Role of Implicit and Explicit Prosody in Sentence Processing

78 Uta Wolfensteller
Habituelle und arbiträre sensomotorische Verknüpfungen im lateralen prämotorischen Kortex des Menschen

79 Päivi Sivonen
Event-related brain activation in speech perception: From sensory to cognitive processes

80 Yun Nan
Music phrase structure perception: the neural basis, the effects of acculturation and musical training

81 Katrin Schulze
Neural Correlates of Working Memory for Verbal and Tonal Stimuli in Nonmusicians and Musicians With and Without Absolute Pitch

82 Korinna Eckstein
Interaktion von Syntax und Prosodie beim Sprachverstehen: Untersuchungen anhand ereigniskorrelierter Hirnpotentiale

83 Florian Th. Siebörger
Funktionelle Neuroanatomie des Textverstehens: Kohärenzbildung bei Witzen und anderen ungewöhnlichen Texten

84 Diana Böttger
Aktivität im Gamma-Frequenzbereich des EEG: Einfluss demographischer Faktoren und kognitiver Korrelate

85 Jörg Bahlmann
Neural correlates of the processing of linear and hierarchical artificial grammar rules: Electrophysiological and neuroimaging studies

86 Jan Zwickel
Specific Interference Effects Between Temporally Overlapping Action and Perception

87 Markus Ullsperger
Functional Neuroanatomy of Performance Monitoring: fMRI, ERP, and Patient Studies

88 Susanne Dietrich
Vom Brüllen zum Wort – MRT-Studien zur kognitiven Verarbeitung emotionaler Vokalisationen

89 Maren Schmidt-Kassow
What's Beat got to do with ist? The Influence of Meter on Syntactic Processing:
ERP Evidence from Healthy and Patient populations

90 Monika Lück
Die Verarbeitung morphologisch komplexer Wörter bei Kindern im Schulalter:
Neurophysiologische Korrelate der Entwicklung

91 Diana P. Szameitat
Perzeption und akustische Eigenschaften von Emotionen in menschlichem Lachen

92 Beate Sabisch
Mechanisms of auditory sentence comprehension in children with specific language impairment and children with developmental dyslexia: A neurophysiological investigation

93 Regine Oberecker
Grammatikverarbeitung im Kindesalter: EKP-Studien zum auditorischen Satzverstehen

94 Şükrü Barış Demiral
Incremental Argument Interpretation in Turkish Sentence Comprehension

95 Henning Holle
The Comprehension of Co-Speech Iconic Gestures: Behavioral, Electrophysiological and Neuroimaging Studies

96 Marcel Braß
Das inferior frontale Kreuzungsareal und seine Rolle bei der kognitiven Kontrolle unseres Verhaltens

97 Anna S. Hasting
Syntax in a blink: Early and automatic processing of syntactic rules as revealed by event-related brain potentials

98 Sebastian Jentschke
Neural Correlates of Processing Syntax in Music and Language – Influences of Development, Musical Training and Language Impairment

99 Amelie Mahlstedt
The Acquisition of Case marking Information as a Cue to Argument Interpretation in German An Electrophysiological Investigation with Pre-school Children