A Perspective Theory of Music Perception and Emotion

Björn Vickhoff

Skrifter från musikvetenskap, Göteborgs universitet nr 90

UNIVERSITY OF GOTHENBURG

Abstract

Doctoral dissertation in musicology at the Department of Culture, Aesthetics and Media, University of Gothenburg, Sweden, 2008, 311 pages. Language: English.
Title: *A Perspective Theory of Music Perception and Emotion.*
Author: Björn Vickhoff
Series: *Skrifter från musikvetenskap, Göteborgs universitet nr 90*

In order to answer the question of why music moves us emotionally we need to understand the perception of music. A perspective theory is suggested. A perspective implies a frame of reference. This means that the perspective selects content – an implicit meaning. This content is an important agent in the perception of form. The general assumption that we judge the content from the form (signification) can thus be challenged. The opposite is also possible: the content makes us perceive form. This is a new understanding of the old question of whether music is a representative or just form. The meaning is immanent in the form.

These processes are pre-attentive. It has been shown that we do not have to perceive music consciously to be emotionally affected by it. This has an interesting consequence. We do not always connect the emotion caused by the music to the music heard. This is obvious in the case of film music. The emotion could attach to anything that we happen to attend. If we pay attention to film characters, the emotion is ascribed to them. If we attend to the own body, the emotion is ascribed to the body (as feelings). Only when we focus on the music the emotion is ascribed to the music and we experience the music as emotional.

The perspectives are discussed in the light of target/source domain theory. A target domain is typically an abstract and complicated behaviour. Source domains are concrete behaviours of survival value. It has been suggested that music is the target domain of several source domains. I argue that the perspectives are source domain processes. Musical emotions are discussed as functions of these source domains: the perception of sound, identification with the other, navigation, and joint action. The perspective theory potentially offers answers to several questions in music philosophy.

Key words: *perception-action theory, emotion, perspective, imagery, enhanced formalism source domains and target domains.*

bjorn.vickhoff@musicology.gu.se

Layout: Mats d Hermansson
Printed by Intellecta DocuSys, Västra Frölunda 2008

ISSN 1654-6261
ISBN 978-91-85974-06-1

Distribution: Department of Culture, Aesthetics and Media
P.O. Box 200, SE-405 30 GOTHENBURG

The body is our primary instrument.
Naná Vasconcelos

Acknowledgements

My supervisors Olle Edström (musicology) and Helge Malmgren (philosophy) invested interest and knowledge.

Collegium SSKKII has been a valuable cross-disciplinary (semiotics, cognition, communication, interaction) reference group.

Martha Ulhua supervised me in Música Popular Brasileira.

The family Gonzaga, Jota Moraes (arranger and keyboardist), and Nelson Faria (my teacher of Brazilian music) opened their homes and hearts in Rio de Janeiro and Belo Horizonte and generously provided me with material and helped me with contacts in Brazil.

Ulrik Volgsten (musicology), Ove Sernhede (culture and psychoanalysis), Elisabeth Ahlsén (neurolinguistics), Åsa Abelin (prosody) and Maj-Lis Follér (Brazilian anthropology) contributed with their expertise.

Every musician I listened to and played with has contributed to this work.

Angus Hawkins proofread the text.

Mats d Hermansson did the layout.

Thank you!

Table of Contents

Introduction

In 1994, I was sitting in a church in my home village Olofstorp, Sweden listening to a speech concerning social work in Brazil. To round up, the speaker showed some tourist pictures from Rio de Janeiro and put on a tape with Brazilian music.

I heard a distinct but humble voice smoothly moving over floating, dynamic guitar chords. It was not a samba or bossa but it still felt Brazilian. For some reason I was deeply moved. The melody was simple and yet perfect, performed in the most sensitive way. I could hear the richness of associations in the performance. The rhythm was delicate. The voice was charged with closeness and sincerity.

On my way home I was convinced that something had been proposed to me and that I had to follow this new trail.

I began to explore what I had heard. The song was 'Asa Branca' performed by the Brazilian singer/composer Luis Gonzaga Junior (1945–1991) at a live concert just days before his tragic death in a car accident. Asa Branca was written by Gonzaguinha's father, Luis Gonzaga, one of the icons of Brazilian music.

The thing is that when I first heard the music I did not know anything about Gonzaguinha, nor did I understand a single word of the song. So, how could this song move me so much?

I managed to get hold of the CD 'Cavaleiro Solitario'[1] and found all the songs wonderful. I played the CD over and over on the car stereo that summer and always got goose flesh on my arms, and tears in my eyes, and when I tried to talk about it with my passengers I was unable to speak. In fact it was ridiculous and quite tiring, but I could not do anything about it. I had to play it, and I always had to go through these emotions.

There was one song 'Fotografia' that shook me more than any other. I made my own epigone version of the music and wrote lyrics in Swedish. Later I found out that the lyrics of the two versions connect emotionally. They are both about sorrow, longing and

[1] Gonzaguinha (1993) CD: *Cavaleiro Solitario*, Som Livre

death. I have to draw the conclusion that this is not merely a coincidence.

Why did I feel addressed by this singer? There was nothing connecting us. He was a mulatto brought up by foster parents in the favela São Carlos in Rio de Janeiro, close to the red light district Estácio – the very cradle of samba music. It was far from my little town Växjö, Sweden where I grew up as the son of the town physician who used to spend an hour after work playing classical music on the grand piano.

The interest of how Gonzaguinha touches me emotionally took over. My need to investigate this was aroused by an interest in my own reactions. I thought I could learn something about myself here. I wanted to cover his shadow with my own body.

On Christmas Eve 2004, some friends had gathered in a neighbour's house to celebrate the holiday. We played the Gonzaguinha song *É* on the CD-player. The guests spontaneously took each other's hands and in the second verse we were all on our feet dancing, still holding hands. I asked them what they thought this song was about. The host's mother immediately shouted out from her wheelchair that it is about the pride of the people. This remark was absolutely accurate. This is what the lyrics are about. How could she tell?

I wanted to write about my emotional reaction to Gonzaguinha, but it was hard to make this a scientific project. There were two obvious dangers with this project: It was about emotions and it was subjective. It had the sign of woolliness written all over it.

I became interested in the general question of how music communicates emotions. My interest was provoked by prevailing semiotic thinking in musicology. The general claim is that music structures are signs of extra-musical meaning. Music was treated as a language with syntax and semantic meanings to be decoded. I saw this theory as too intellectualistic. It did not fit with my experience of working with mentally retarded children suffering from Down's syndrome. These children are very affected by music. It did not fit with my experience as a father, since music was the first way of communication with our kids. It did not fit with my experience as a professional musician. It disturbed me that music was explained as an abstraction of something else. The beautiful experience was reduced to intellectual understanding.

I have realized that *perception-action theory* could offer a more body oriented and direct explanation of musical emotions. An article on empathy (Preston & de Waal 2002)[2] gave me the idea that mirror neurons could have something to do with musical emotions. In 2004 Helge Malmgren and I wrote an article suggesting this possibility (Vickhoff & Malmgren 2005)[3]. This article is presented in a revised version as Chapter 3 of this dissertation. Mirror neuron reactions imply that we take the perspective of the other. This fact made me start thinking about other possible perspectives in music listening and what they would imply in terms of perception and emotion.

Problem formulation

The problem with music and emotion is not that music may make us think of something that might move us emotionally. If music signifies something delightful or something terrible, this will affect us. The problem is that music seems to move us emotionally even if there is no such obvious reason. This is the enigma.

In 1979 Philip Tagg wrote *50 seconds of television music – towards an analysis of affect in popular music.* This is an influential dissertation on music and emotion combining *semiotics* (the study of signs) with communication theory. In the conclusive chapter Tagg expresses the difficulty in explaining the music emotion phenomenon in words:

> If we were to analyze or explain in cognitive verbal terms what exactly happens in another type of nonverbal communication, for example when someone strokes the cheek of a loved one and

[2] Preston, S.D. & de Waal, F.B.M. (2002) Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences.* Vol. 25. Issue 1. 1–72

[3] Vickhoff, B. & Malmgren, H. (2005) Why does music move us? *Philosophical Communications.* Web Series, No. 34. Dept. of Philosophy, Göteborg University, Sweden. ISSN 1652–0459

> looks longingly into his/her eyes we would find ourselves up to a similar problem. (Tagg 1979: 231)[4]

This parallel is accurate. It comes closer, I would say, to the listening situation that reading a book does (decoding signs). Thus it points rather to a phenomenological approach than to semiotics. This hunch is the starting point of this dissertation.

Phenomenological theory is focussed on the direct subjective experience of the present moment. Progress in perception-action theory (which overlaps with phenomenology) and new findings in brain research supporting perception action theory, have made it possible to analyse the music-emotion question from new perspectives. The focus is moved from explicit understanding to implicit understanding. New methods, such as ERP (event related potential) and brain scanning, allow us to study the perception of music on lower levels of consciousness. These studies indicate that we do not always, in everyday listening, analyse what we hear, but perceive it directly. We live in the movement of the music, rather than take it to be a sign of something. In fact music can be perceived without being attended to. What does this tell us about the music emotion question?

Question

The question propelling this dissertation is: How are we moved emotionally by music? This question can be broken down into three questions:

1. How do we perceive music?
2. What is the connection between music perception and emotion?
3. What is the nature/function of music elicited emotions?

[4] Tagg, P. (1979) *50 Seconds of Television Music – towards an analysis of affect in popular music.* Studies from the Department of Musicology. No 2. Gothenburg

As can be seen from these questions, this subject entails major musicological questions. To answer these questions other issues have to be discussed concerning the understanding of the concepts emotion and perception, the relation between emotion and perception, as well as different levels of consciousness of emotions and perceptions.

We cannot avoid questions such as: What is the nature and function of music elicited emotions? Are they of the same kind as everyday emotions? And if so, what everyday emotion should they be compared to? Is a music-elicited emotion a reaction to an object? Is it a reaction to a representation of an object? Is it a reaction to a form? Is it a reaction to something imagined? Is it an emotion communicated? Is it a reaction to expectancies and the fulfilment/violation of these expectancies? Or is it something else?

Background and previous research

In the following, we will scan previous research and theories on music and emotion. The field is cross-disciplinary and for this reason the terminology is not coherent. The purpose is to inform the reader about the diversity of approaches and present the main lines in the development of these theories. In the chapters to come we will have many occasions to reconnect to these theories.

The question of why music affects us emotionally has attracted philosophers. It can be spotted already in Plato's dialogue with Socrates in *The State,* where the subject of emotional properties of scales is discussed.

Philosophers have a way of asking seemingly naïve questions. In this case the question is: Where is the emotion? Is it in the performer, in the music or in the listener? As we will see, the theories on music and emotion reflect this question.

We can see a pendulum movement between two positions in the view of music. One holds that music is *representative* and the other that music is pure *form.* If music is representative there are two options: it could represent emotions or it could represent an object that we can react to emotionally. On the other hand, if music is just form it is hard to see why it moves us emotionally at all.

During the Baroque era the *Figurenlehre* was developed (Wessely 1972)[5]. This was a composer's dictionary of music emotions. Rhetoric figures or gestures in music were thought of as representations of emotion in the Baroque opera. The idea was theoretically developed in the *Affectenlehre* (Mattheson 1739)[6]. Jean-Jacques Rousseau rather took a formalist stand in his 1761 article *Essay on the origin of languages*. Rousseau's *mimetic theory* drew on resemblances between emotional expression in music and emotional expression in speech (Rousseau 1986)[7].

Immanuel Kant discussed the concept *aesthetics* 1790 in the second part of *Critic of Judgement* called *The Critique of Aesthetic Judgement* (Kant 1952)[8]. Kant did not restrict aesthetics to art – anything could give an aesthetic experience (Kivy 2002)[9]. Kant claimed that aesthetic judgement is based on form and content. Our personal feelings and beliefs are affected by content, whereas a universal judgement of beauty is based on form. We contemplate form in a "disinterested" way – in a free play of imagination and knowledge. Kant's ideas inspired the musician and critique Eduard Hanslick.

Hanslick was the first person to systematically discuss form and content in music. His *Vom Musikalisch Schönen* (On the Musically Beautiful) from 1854 contains a critique of Mattheson's *Affectenlehre* (Hanslick 1986)[10]. Hanslick arrived at the conclusion that music cannot evoke emotions. He presented three main arguments:

1. "The content of music is tonally moving forms" (p. 29). In the music's themes, "form and content cannot be separated" (p. 81). Music is made of tones – tones are its content, not to be confused with the "subject matter" inspiring it (p. 78). To have a subject matter as content, the extra-musical element

---

[5] Wessely, O. (1972) *Musik.* Habel, Darmstadt

[6] Mattheson, J. (1739) *Der volkommene Capellmeister.* Herold, Hamburg

[7] Rousseau, J.J. (1986) Essay on the origin of languages. *On the Origin of Language: two essays.* (Eds. J.H. Moran & A. Gode) University of Chicago Press, Chicago. 5–74

[8] Kant, I. (1952) *The Critique of Aesthetic Judgement.* Clarendon, Oxford

[9] Kivy, P. (2002) *Introduction to a Philosophy of Music.* Clarendon, Oxford

[10] Hanslick, E. (1986) *On the Musically Beautiful.* Hacket Publ., Indianapolis

must be definite and we must be able to formulate it in words. This is not the case as we have no agreement on the "definite" description to give it. So, it can only have an "indefinite" content. Since music cannot represent extra-musical objects it cannot represent objects that we react to emotionally.

2. There is no agreement between listeners about what emotion they hear in music.
3. It is unthinkable that music could be dispositionally emotional, meaning that it could not upset you the way an event (such as an accident) can.

Hanslick was thus a formalist. The prevailing understanding of emotions at this time was Kantian. To Kant "emotion" implied an object. We do not just love or hate; we love or hate something. There must be an object for the emotion. Since this object cannot be represented in music, Hanslick argued, music cannot cause emotional reactions.

Hanslick's arguments have been contested. In *Emotion and meaning in music* Leonard Meyer dismisses the argument that emotions must have an object as a confusion of expressionism with referentialism (Meyer 1956)[11]. As we will see, emotion is a concept that has a tendency to escape definitions. All definitions of emotion do not require consciousness about an object. Hanslick's second point is overthrown by numerous experiments in music psychology. There is a significant correlation between listeners concerning what emotion they perceive in a piece of music. However, Hanslick's idea that the content of music is the sounding and moving form is interesting. In this respect, he was ahead of his time. He argued that content was not represented by the form, but is inseparable from it. This theme will be developed in Chapter 1.

The 20th century witnesses interesting ideas about music and emotion. The psychologist Carroll C. Pratt claimed that "[m]usic can be agitated, restless, triumphant, or calm since it can possess the character of the bodily movements which are involved in the moods and

---

[11] Meyer, L.B. (1956) *Emotion and Meaning in Music*. University of Chicago Press, Chicago

emotions that are given these names" (Pratt 1931: 7)[12]. This leads to his much-quoted sentence: "Music sounds the way emotions feel" (p. 203). As far as I know, this is the first time anyone made the connection between musical emotions and body movements. As it turned out, this idea would become a major theme.

Pratt inspired Susan Langer who argued in *Philosophy in a new key* that music does not express, but merely mimics the time course of emotional experience (Langer 1942)[13]. With this we are back to representation. Langer used the term *isomorphism* (structural likeness). This likeness caused Langer to propose that music generally could be seen as an icon of emotional life, but not specifically that a certain piece of music could possess a certain feeling. In *Feeling and Form* Langer proceeded by claiming that music is a symbol of the particular feeling domain of *transience* (the experience of the passing of time) (Langer 1953)[14]. Langer saw tonally moving forms as iconic of the pattern of time. Music thus in her view presents "virtual" time – a directly felt audible illusion or image of psychological time.

Philip Alperson (2004)[15], among others, questioned Langer's view that music expresses *ideas* (iconically and symbolically) rather than definite *emotions*. Langer's distinction between musical and psychological time was criticized. The critique can be summarized as a questioning of if listening to music "in effect is presented with ideas about something else" such as the life of feelings.

We can see that the seemingly naïve question concerning the emotions whereabouts has turned out to propel theory development. I will come back to a list of the currently most influential theories, but first I will briefly present the current state of the research. Let me start this presentation with an observation by two neurophysiologists:

---

[12] Pratt, C. (1931) *Emotion and Meaning in Music: a study of psychological aesthetics.* McGraw-Hill Book Company Inc., New York and London

[13] Langer, S. (1942) *Philosophy in a New Key: a study in the symbolism of reason, rite, and art.* Harvard University Press, Cambridge, MA

[14] Langer, S. (1953) *Feeling and Form.* Charles Scribner's Sons, New York

[15] Alperson, P. (2004) Music: formalism and beyond. *The Blackwell Guide to Aesthetics.* (Ed. P. Kivy) Blackwell Publ., Malden, MA. 254–275

> It is remarkable that any medium could so readily evoke all the basic emotions of our brains (and much more), coaxing us to consider our innermost nature and to savour the affective dimensions of our minds. (Panksepp & Bernatzky 2002)[16].

The present knowledge in the field is mirrored in the anthology *Music and Emotion – theory and research* (Juslin & Sloboda 2001)[17]. Contemporary research on music and emotion is also summarized in the overview *Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening* (Juslin & Laukka 2004)[18].

Research within music psychology unambiguously shows that music arouses emotions. This has been studied in physiological studies of ANS (Autonomic Nervous System) reactions (such as electro dermal response, blood pressure, breathing etc.) as well as in verbal self-reports. There is a correlation of up to 0.80 between subjects on what emotion they identify in an instrumental piece of music. The question now is not *if*, but *how* music affects us. Although Juslin's and Sloboda's book covers an enormous body of research, this question is far from solved. One indicator of this is the disciplinary span of the theories. The contributing authors represent philosophy, musicology, psychology, neurology, anthropology, sociology, music therapy, film music, communication and medicine. The reason for this diversity of attempts is that emotion is a concept that can be discussed at all levels. So can music.

I would like to emphasize one paper because it has bearing on the framework of this dissertation. In *Strong experiences related to music: a descriptive system* by Gabrielsson & Lindström Wik (2003)[19], listeners were asked to write about their strongest music experi-

---

16 Panksepp, J. & Bernatzky, G. (2002) Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes.*Vol. 60. No. 2. 133–155

17 Juslin, P.N. & Sloboda, J.A. (2001) *Music and Emotion – theory and research.* Oxford University Press, Oxford

18 Juslin, P.N. & Laukka, P. (2004) Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research.* Vol. 33. No 3. 217–238

19 Gabrielsson, A. & Lindström Wik, S. (2003) Strong experiences related to music: a descriptive system. *Musicae Scientiae.* Vol. 7. No 2. 157–217

ences. The conclusion of this article is that strong experiences depend on three factors: the music, the listener, and the situation.

Objective

The objective of this work is to present a coherent theory of the perception and emotion of music.

Theory

This work relies on perception-action theory. This theory has been developed by the French philosopher Maurice Merleau-Ponty (1907–1961) and his followers. I will present this theory and develop a perspective model for music in Chapter 2.

Material

As mentioned the music-emotion question is cross-disciplinary. The material used here is scientific reports from the fields of neurophysiology, animal navigation, psychology, musicology, communication, anthropology, and perception-action philosophy.

I will exemplify theoretical statements with everyday observations. These examples are primarily taken from my experience as a professional musician[20].

Delimitations

This is to a certain extent a subjective investigation. Subjectivity follows from the perception-action approach. As I intend to show, every perspective of perception entails the integration of an egocentric perspective. We cannot perceive without being subjective. This

---

[20] *International Who is Who in Music* (annual) Vol 2. Popular Music. (Ed. S. Tyler) Melrose Press Ltd., Cambridge, England

is indicated in the concept *intersubjectivity.* Some see subjectivity as a limitation, or worse – as unscientific. But I believe that it is impossible to write about certain topics without using a subjective experience. Music and emotion is one of them. Thus I see it as an asset that I have a personal lived experience of writing, producing, arranging and performing music in ensembles as an artist and as a bandleader. I will use this. This will however be balanced by references.

The ambition is limited to presenting a theory of music perception and emotion. This does not mean that I claim to answer the music/emotion question. I do not intend to prove the theory, just to propose it and present theoretical arguments, evidence and everyday observations to support it. Some of the hypotheses underpinning this theory are supported by evidence whereas others are not. They just follow from the context they are a part of. And in some cases they follow from the coherence of the theory.

The work has the limitations and advantages of cross-disciplinary research. I have tried to avoid mistakes through discussions with experts. One such forum has been SSKKII, a cross-disciplinary group where scientists gather to discuss topics of mutual interest concerning language (in every aspect), semiotics, communication, cognition, information and interaction. This group has gathered scholars from the fields of linguistics, philosophy, psychology, acoustics, cognition and musicology. Other occasions for testing ideas have been Ph.D. courses and symposia.

Since this work to a considerable extent relies on neurological evidence, we may ask what this type of evidence can tell us about the perception of music. Brain scanning enables us to take a glimpse into an area previously considered a black box. Now we do not have to settle with the documentation of the reactions to stimuli. We can actually look into the brain to see what happens when the brain reacts to the stimuli.

But: We do not understand how music is perceived and felt just because we know what brain areas are involved in the process. Even if we knew everything that happens in the brain as we listen to music, the experience of music listening cannot be mapped. It is a vast step from the interpretation of images of the brain as it is exposed to music to the phenomenal experience of music listening. What biology can do is to provide guidelines. Biology can be used to understand pre-attentive processes – to peep behind the screen of consciousness. We need to do this, since, as I intend to show, such processes produce musical emotions. This is where this work becomes cross-disciplinary. The domain of philosophy and particularly phenomenology is traditionally limited to the meaning constituted by consciousness. But, the meaning of a stimulus, in the sense that we respond adequately, can and does arise on pre-attentive levels. For these reasons I will present a taxonomy of different levels of consciousness and relate the discussion to these levels.

Perception is not restricted to certain brain areas but is a process that involves the whole body. The fact that a specific area is activated in a perception can have a number of explanations: The area can be supporting or inhibitory; it can prepare other parts for processing, it can just be a relay station, it can add some necessary but overlooked piece of information, or it can be activated by something else than the stimulus under investigation. Processes in the brain are not exclusively located to specific areas but different areas contribute to the process. Instead we should talk about modules or circuits. Even so there is a risk of oversimplifications as these circuits are integrated.

Contrary to most other activities, music listening activates the whole brain. This indicates the complexity of music. It is a web of rhythm, harmony, timbre and melody, and it is connected to rich associations. It elicits motor activity (Haueisen & Knoesche 2001)[21], visual sensations (Körlin 2005)[22], bodily sensations (Gabrielsson & Lindström Wik 2003) and emotions (Gabrielsson & Lindström 2001)[23]. It has the power to engage the listener at all levels.

---

[21] Haueisen, J. & Knoesche, T.R. (2001) Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience.* Vol. 13. No 6. 786–792

[22] Körlin, D. (2005) *Creative Arts Therapies in Psychiatric Treatment – a clinical application of the Bonny Method of guided imagery and music (BMGIM) and creative arts groups.* Doctoral dissertation. Karolinska Institutet, Stockholm

[23] Gabrielsson, A. & Lindström, E. (2001) The influence of musical structure on emotional expression. *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press, Oxford. 223–248

The brain is not, as previously assumed, like a Swiss army knife, where every part has its separate function (Quartz & Sejnowski 2002)[24]. On the contrary, it should be compared to an orchestra where all the sections conjoin. This calls for holistic approaches and top-down understandings. The functions of the brain areas involved in music processing are not clear. Although there are areas proven vital to music processing, there simply is no such thing as a "centre of music" somewhere in the brain. If a vital part is damaged and the ability to perceive music is lost, this just indicates that a link in the process is broken.

The technique to sort out what is caused by a certain stimulus is called masking. A control picture is taken under test conditions with the stimulus to be tested excluded. This picture is used to mask irrelevant activity. Knowledge from research on the consequences of brain lesions as well as surgery and electrical stimulation of brain areas help interpreting the images.

In addition, the *plasticity* of the brain has to be addressed. Plasticity means that the brain is in a constant process of adapting its functions to the environment. Living in the world literally helps build our brain (Qartz & Sejnowski 2002). The brain of the musician exhibits several examples of this (Weinberger 2004)[25]. Training affects enlarges areas for tone detection and sensitivity to the sound of the instrument, coordination of hands, motor skills and the sensory and motor representation of the fingers used when playing.

I think it is helpful to keep these reservations in mind, as the interpretation of the presentation of findings otherwise may be too simplistic. Any finding in this complex area must be put in relation to a vast context of theory and findings.

---

[24] Quartz, S. & Sejnowski, T. (2002) *Liars, Lovers, and Heroes – what the new brain science reveals about how we become what we are.* HarperCollins Publishers, New York

[25] Weinberger, N.M. (2004) Music and the brain. *Scientific American.* Nov. 89–95

Theories on music and emotion

The multidisciplinary character of the research of music and emotion is striking. These different levels, spanning from brain physiology to culture, have produced a range of theories. In these theories we can always spot a stance in the old philosophical question of whether music is representative or just form.

Expectation theory

Leonard Meyer can be described as a formalist in the sense that he discusses how the structures of music may affect us emotionally. The general idea is that the perception of music structure follows Gestalt laws (e.g. law on good continuation, law on closure). These laws generate expectancies. Emotions are generated by hindrance or fulfilment of these expectations (Meyer 1956). The Gestalt approach is fruitful. We will come back to this in Chapter 1 and 4.

A new version of expectation theory is presented by David Huron in *Sweet anticipation* (2006)[26], where Huron draws on new findings in the psychology of expectation. These findings appoint expectations to be derived from learning (exposure) rather than from Gestalt. The book presents a theory of music and emotion grounded in the effect of musical surprise (deviation from the expectation). Huron's theory will be discussed in Chapter 4.

The semiotic approach

The semiotic approach is grounded in the assumption that music represents something – signifies something. Semiotics became widely used in anthropology and made its way into musicology via music anthropology in the 70s. Music was studied as a sign (artefact) of culture. From there on music semiotics took the huge leap to claim that music structures are semiotic signs of affective meaning (Tagg 1979). Tagg argued that these signs developed during the course of

---

[26] Huron, D. (2006) *Sweet Anticipation: music and the psychology of expectation.* MIT Press, Cambridge, Massachusetts

music history. He adopted the concept *museme*, from Charles Seeger. It corresponds to the linguistic concept morpheme. The museme is defined as “the basic unit of musical expression which in the framework of one given musical system is not further divisible without destruction of meaning” (p. 71). These smallest units of meaning could be defined by changing a note or two to see at what point the meaning is changed. This approach can be traced to the linguist Louis Hjelmslev’s commutation principle (Hjelmslev 1953)[27]. It was argued that the composer uses musemes to create affective meaning. In other words affective meaning is encoded as musemes and the listener is decoding these signs. This makes music a language.

Music semiotics has developed into a vast and diversified field. All kinds of semiotic possibilities have been explored: symbols, indexes, icons, structural approaches, gestures, syntactical approaches etc. For an overview see Monelle (1992)[28].

The semiotic approach could be criticized for being a somewhat unmusical bottom up theory where the meaning is derived from its entities – the musemes. As noted already by Christian von Ehrenfels in his paper *Über 'Gestaltqualitäten'* 1890 the perception of a melody cannot be derived from its single isolated notes but from a succession of notes combined in a meaningful way (Ehrenfels 1988)[29].

The Brunswikian lens model

Brunswik’s model (from 1956) is originally a model of perception that has specifically been applied to the communication of emotion in *prosody* (the sound aspect of language). The model has been adapted to emotion in music (Juslin 2001)[30]. The performer is

[27] Hjelmslev, L. (1953) *Prolegomena to a Theory of Language*. Baltimore: Indiana University Massachusetts

[28] Monelle, R. (1992) *Linguistics and Semiotics in Music*. Harwood Academic Publishers, Edinburg

[29] Ehrenfels, C. von (1988) *Foundations of Gestalt Theory*. (Ed. and translation B. Smith) Philosophia, Munich and Vienna. 82–117

[30] Juslin, P.N. (2001) Communicating emotion in music performance: A Review and theoretical framework, in *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press, Oxford. 309–337

thought to encode her emotions into cues – bits of information (such as volume, timbre, and tempo) making up the music structure. These signs pass through an information channel and are then decoded by the listener who is judging the emotion. Even if each cue does not discriminate significantly between emotions, the cues sum up to a pattern indicating a certain emotion with significant probability – the larger the number of cues, the more reliable the prediction. Scherer (1986)[31] observed that listeners have found it difficult to identify the acoustic cues behind their predictions. This is the reason for studying patterns of cues rather than specific cues. The model does not account for how these signs have developed their emotional meaning or where/how this decoding takes place.

Arousalism

Arousalism states that art properties evoke emotions as good or terrible news or events would arouse us (Matravers 1998)[32]. This theory locates the emotion in the listener rather than in the music. The argument is that music cannot possess emotions since music is not a sentient being. For this reason the emotion must be aroused in the listener. Arousalism has been criticised since it does not consider expressiveness as a music property, but confuses it with the emotions of the listener. It does not explain why this music makes us happy while that music makes us sad.

Persona theory

Persona theory connects music expressiveness to some imagined person. As we have already seen Carroll Pratt and Jean-Jacque Rosseau pointed in this direction. The persona theory was developed by Walton (1988)[33] and it has many followers. The so called "strong

---

[31] Scherer, K.R. (1986) The vocal affect expression: A review and a model for future research. *Psychological Bulletin*. Vol. 99. No 2. 143–165

[32] Matravers, D. (1998) *Art and Emotion.* Clarendon, Oxford

[33] Walton, K.L. (1988) What is abstract about the art of music? *Journal of Aesthetics and Art Criticism*. Vol. 46. No 3. 351–64

position" is taken by Levinson (1996)[34] who defines a musical passage as expressive, if and only if (as philosophers tend to say) it is heard as an expression of that emotion by an imagined human subject, the music's persona. Levinson argues that the listener feels sympathy (shares her emotions) with the persona, but that this persona in fact is a projection of the listener, resulting in sympathy for ones own feelings. This makes Levinson open to the same criticism as against arousalism – his theory does not explain the expressiveness of the music.

Persona theory is an example of the perception of music as *the other*. "The other" is a concept in psychoanalysis as well as in philosophy. In this thesis "the other" refers to the movements/expression of another human being. We will discuss this in Chapter 3.

Emotional contagion

Emotional contagion is described as "the tendency to automatically mimic and synchronize facial expressions, vocalizations postures and movements with those of another person and, consequently, to converge emotionally" (Hatfield, Cacioppo, & Rapson 1993: 5)[35]. The emotion is *induced* directly and pre-attentively in the listener by expressive performances (Juslin 2001). It is likely that we get aroused by the voice-like aspects of musical performances.

This theory implies a low level of consciousness. The emotion is felt but not thought of. Different levels of consciousness will be discussed in Chapter 1.

Metaphor theory

In his essay *Understanding Music* (1983)[36] philosopher Roger Scruton points out a crucial distinction between music and environmental

---

[34] Levison, J. (1996) *The Pleasures of Aesthetics*. Cornell University Press. Ithaca, NY

[35] Hatfield, E., Cacioppo, J.T. & Rapson, R. L. (1993) *Emotional Contagion*. Cambridge University Press, New York

[36] Scruton, R. (1983) Understanding music, *Ratio*. Vol. 25, No. 2. 97–120

sounds. While sounds are material facts, music is an intentional construct – a matter of the concepts through which we perceive the world. This is supported, Scruton argues, by the way music is talked about and analysed in terms of metaphors. Tone is the product of imagination, not material but an idea (Scruton 1997)[37].

Walton's and Levinson's persona theory could be seen as metaphoric. Metaphor theory is used by Nicholas Cook (1990)[38] as a tool for music analysis and for a wide approach of discussing music and meaning. Zbikowski (1998)[39] drawing on Lakoff & Johnson (1980)[40], points out that metaphorical thinking is the structural bases of all cognitive activity. Lakoff & Johnson depict metaphor as a basic structure of understanding through which we conceptualize a complex and abstract *target domain* in terms of a familiar and concrete *source domain*. This is called *cross-domain mapping*. We refer our thinking to bodily experiences. Such a metaphor is called a conceptual metaphor to be distinguished from a linguistic metaphor, which is the verbal expression of a conceptual metaphor. In music a conceptual metaphor could be the map of sounds in space relative to the body. The corresponding linguistic metaphors are for example "high" and "low" pitch. Thinking about the underlying mechanism for cross-domain mapping, Johnson proposed the *image schema*, which is not visual but rather something with the abstract structure of an image.

The importance of this theory is that it provides music with meaning. This theory can explain how music might produce the illusion of concrete reality. In other words cross-domain mapping might produce the object lost in Hanslick's formalism. Although music is not a language that can refer to the external world, it could create illusions of it.

The distinction between sound and music is important. Animals hear the same sounds but do they hear music? Probably not. There

---

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

38 Cook, N. (1990) *Music, Imagination, and Culture*. Clarendon, Oxford

39 Zbikowski, L. (1998) Metaphor and music theory: reflections from cognitive science. *The Online Journal of the Society of Music Theory*. Vol. 4. No 1

40 Lakoff, G. & Johnson, M. (1980) *Metaphors We Live By*. University of Chicago Press, Chicago

are other explanations of how musical meaning is created than cross-domain mapping. We will look into this in the theory chapter.

Cross-domain mapping is nevertheless an intriguing idea. I prefer to understand it in a more biological and less conceptualized sense: The target domain is a complex behaviour that cannot be explained in evolutionary Darwinian terms. This domain is based on source domains that do have concrete survival value. This might explain why we have music. We "just" have to find the source domains that underlie musical behaviour. This will be discussed in the concluding chapter.

Trivedi's conceptions

The idea of the conceptual metaphor brings us to the music philosopher Saam Trivedi. Trivedi develops the idea that music touches us because it triggers our fantasy to make conceptions (Trivedi 2003)[41]. Fantasy is propelled by iconic resemblances between musical sounds and sounds in our everyday environment, by resemblances between musical moods and our own state of mind, or by resemblances between music and human expressions. These conceptions are seen as perceptual or cognitive illusions. Music can resemble natural sounds, which we identify as sound sources and thus as objects. These objects can carry emotional value. Trivedi's last point (music resembles human expressions) has an affinity with the persona theory. Trivedi makes the distinction that music does not resemble humans. What it resembles is rather the temporal structure of human emotional expression. This idea comes close to Daniel Stern's *vitality affect* (Stern 1985). Stern discusses the temporal contour of emotions, and he uses musical terminology to describe this (fading, accelerando, exploding etc.) Trivedi unites iconicity, arousalism and persona theory and treats them all as make-belief conceptions.

[41] Trivedi, S. (2003) The funerary sadness of Mahler´s music. *Imagination, Philosophy and the Arts.* (Eds. M. Kieran & D. McIver Lopes)Routledge, London. 259–271

Enhanced formalism

I will give the philosopher Peter Kivy extra space. The reason is that I, as a musician, like his so-called *enhanced formalism*. Apart from that, Kivy is provocative and fun to read although some of his ideas must be completely wrong.

Kivy lustfully kills all semiotic claims of narration in music. In *Introduction to a Philosophy of Music* (2002) Kivy discusses the repetition element in music. This element would be impossible in a narrative. "Suppose", Kivy argues, "Hamlet was constructed that way. Then, instead of saying *To be, or not to be*…once, and then get on with his life, Hamlet would repeat every five minutes, *To be, or not to be*…" (p. 154). This indicates that music and narrative fiction are completely different art forms. Kivy criticizes every attempt to discuss narration in music and exemplifies with Susan McClary's texts on narratives of gender in classical music. If, Kivy asks, this is the meaning of the piece, why then is she alone of that opinion? Thus, music cannot represent. The emotion in music must be perceived in the form.

To Kivy music is form, although in an enhanced sense, which includes perceivable properties and qualities. One of these properties inscribed in the musical form is emotion. Emotion can be detected in the music pattern – or movement, just as it can be detected in human movement patterns. As we can see this comes close to Pratt's and Langer's view that music can possess the character of the body movements reflecting moods and emotions. In contrast to Langer though, Kivy does not treat music structure as iconic signs of emotions. Rather he denies that music is about something else. Music has emotions "as a perceptual part of its structure". The emotion is *in the contour*. We do not, as Walton or Levinson, have to imagine a persona to be affected. We are just used to reacting to this kind of *temporal topography* (Kivy 1989).

Now let me exemplify where I believe Kivy goes wrong. There is something, he says, in sad music that is perceived as sad. This is not because the composer was sad when he wrote the piece and it does not arouse sadness in the listener. Kivy exemplifies with the sad look in the face of a dog. The dog was born that way. The sad look is not telling anything about the emotion of the dog. It *appears* that way. Neither does this sad face affect us, because we know that as long as this creature is wagging the tail, she is happy. This expression-appearance theory has been quite influential (Davies 2001: 32)[42]. In Kivy's view we perceive emotion in music, but we are not aroused by it. This dubious reasoning paints Kivy into a corner, since he cannot deny that music affects us. It does affect us, he says, but for other reasons. The unfortunate result of this is that the promising approach called enhanced formalism loses explanation value. Kivy suddenly turns to another explanation: The old *object – belief – feeling* formula. Although music cannot represent an object, it is an object – an artefact – in itself. And it is an object for human beliefs. If we believe that the music is magnificent and possesses high aesthetical values we get a feeling of excitement and awe.

I disagree. For one, we cannot rule out that the emotional state of the composer does affect the composition. Secondly, it is not self evident that the emotions we get from musical listening are based in *object-belief-feeling*. We can be moved by the perception anyway. Thus we do not need the explanation that the music is of high aesthetic value. Gabrielsson and Lindström Wik (2003) have shown that any music could produce awe. And: It is not self evident that emotion from music is based on reason. Quite on the contrary the belief that the music is magnificent could to a large extent be based on the felt emotion. Kivy admits that some people can possess emotions that are not based in reason, but he dismisses them as "odd, some of them inexplicable, or downright abnormal" (Kivy 2002: 126). And so it seems, when people show emotional reactions for no obvious reason. However reason is not a condition for emotions. If somebody is laughing, this makes me laugh. Automatically. I do not need a reason for it. Laughter can be contagious.

Kivy is putting the listener in a dependent situation. Judgements of aesthetic value (such as stating the magnificence of a certain piece) are the judgements of the elite. It is the canon. Not until we believe in that judgement (and make ourselves part of that elite) are we allowed to be touched. This cannot be true. I think Kivy's idea about

---

[42] Davies, S. (2001) Philosophical perspectives on music's expressiveness. In *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press. Oxford. 23–44

magnificent music is a long shot to provide an object for the emotion.

Aside from these objections it is clear that we are able to perceive the emotion of somebody from facial expression, tone of voice and body movement patterns. I will argue that we do this without reason. When we are involved in a conversation, we focus on the content of the message and are less conscious of the facial expression, the tone of voice, or the gestures of the other. These parameters are usually not attended to. Yet these expressions affect us. They automatically affect our own feelings, our own tone of voice, our own facial expression and our own gestures. And what is more, they affect our understanding of the message. Kivy's important contribution is the notion that the emotion is expressed by the temporal topography. In other words emotions can be perceived from the form. This does not imply that we have to make up a persona. This is right. We may very well listen to music without making up some imagined person; we may just react emotionally on the movement patterns the same way as we react on the bodily movement patterns of another being.

Beside these theories I ought to mention the anthropologists Ellen Dissanayake (1992)[43] and (2000b)[44] and Judith Becker (2001)[45]. Both hint at interesting new ways to approach the question. We will return to the anthropological approaches in Chapter 5.

Thesis overview

This thesis is a music-philosophical discussion where I present a *perspective theory of music perception and emotion*. Each perspective in

[43] Dissanayake, E. (1992) *Homo Aestheticus: where art comes from and why.* Free Press, New York

[44] Dissanayake, E. (2000b) *Art and Intimacy: how the arts began.* University of Washington Press, Washington

[45] Becker, J. (2001) Anthoprological perspectives on music and emotion. In *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press, Oxford. 135–161

this model has a specified landmark and a specified frame of reference. A perspective is thought of as a situation adapted program for neuronal music processing. Each program implies specific emotional reactions.

Chapter 1 *From Soundwave to Music* presents perception-action theory and outlines the implications for music perception. Concepts such as the levels of consciousness, form & content, figure & ground, time & movement, imagery and perspective are discussed. This motivates the perspective theory. From there on the theory, which entails five principal perspectives of music perception, structures the dissertation. The implications for music perception and emotion of these perspectives will be discussed in chapters 2–6 called *Sound, Listening to the Other, Navigating in the Soundscape, Tribal Rites,* and *Contemplating Music.*

In each of these chapters I intend to discuss three overall questions:

1. How is music perceived and processed in this perspective?
2. How does perception connect to emotion in this perspective?
3. What kind of emotion is connected to this perspective?

In Chapter 7 the present state of research on emotions will be presented. We will discuss the connection between perception and emotion generally and compare this with what we have found out so far about music and emotion. I will make some observations that separate musical emotions from other emotions. This will lead to a hypothesis concerning musical emotions.

The perspective theory of music perception and emotion will be summed up in the concluding chapter. The perspectives will be connected to source domains, in order to discuss possible use aspects of musical emotions. The suggested perspective theory will be related to the theories presented in this introduction.

1

From Sound Wave to Music

Towards a Perspective Theory of Music Perception

Merleau-Ponty presented in 1945 the basis for perception-action theory in *Phénoménologie de la Perception* (Phenomenology of Perception, Merleau-Ponty 1996)[46]. Perception-action is not just about how perception leads to action but also about how action leads to perception. I will introduce some relevant ideas of this theory and apply them to music. Additionally I will introduce new brain science evidence that sheds light over this theory.

This chapter aims to propose and motivate a *perspective theory of music perception*. I argue that the perspective steers the choice of content and that this selected content affects the perception of form. Thus the perception of form always entails meaning. Meaning thus does not have to be *represented* by form, but is perceived as an implicit understanding of the form. "Vision is", as stated by Merleau-Ponty, "already inhabited by a meaning (French: sens) which gives it a function in the spectacle of the world and in our existence" (p. 52). This is true for sounds as well. Meaning is implemented already in the perception of sound and of course even more so in the perception of music. The sensory input cannot be consciously perceived. It is the perception process that produces the final perception. This process involves memories. This has direct relevance to the philosophical issue whether music is representative or just form.

[46] Merleau-Ponty, M. (1996) *Phenomenology of Perception*. Routledge & Kegan Paul, London

In this chapter I will introduce a theoretical understanding of implicated concepts: *perception, meaning, figure and ground, form and content, procedural knowledge, and imagery.*

Levels of consciousness

The appraisal theory of emotion claims that emotional states result from appraisal judgements (judgements "to the effect that one is facing a predicament that matters") of a perceived object or situation (Arnold 1960: 171)[47].

In this theory perception is considered conscious. The so called "appraisal dimensions" listed by Richard Lazarus, one of the architects of the theory, are: goal relevance, goal congruence, type of ego involvement, blame or credit, coping potential and future expectancy (Lazarus 1991)[48]. Thus, the theory states that an observer becomes aware of an object and explicitly judges these dimensions. An appropriate emotion is supposed to result. Is this so?

Jesse J. Prinz takes a different stand as he claims that the judgements are implicit "gut reactions" (Prinz 2004)[49]. Is it irrelevant whether we are aware of the music or not? This would be a rash statement. At least it is safe to say that musical associations, since they depend on memories, evoke emotion. Music obviously evokes strong emotions when associated with our nation, first love, home, childhood, parents etc. Can this association be implicit? That is, can the national anthem, say, touch the listener because it pre-attentively actualizes implicit memories of the nation?

At this stage let us just agree that there is good reason to discuss how music is perceived before we attend to the question, how, and on what levels of awareness, music affects us emotionally.

---

[47] Arnold, M. (1960) *Emotion and Personality.* Columbia University Press, New York

[48] Lazarus, R. (1991) *Emotion and Adaptation.* Oxford University Press, New York

[49] Prinz, J.J. (2004) *Gut Reactions – a perceptual theory of emotion.* Oxford University Press, New York

To discuss perception we need to distinguish levels of consciousness. This is a disputed topic. I am going to use the following principal levels:

1. *Cognitive consciousness* referring to the consciousness of facts. It is coded information that can be communicated directly or via media. This information can be *unattended* (e.g. we know that Sweden is a Scandinavian country but we are not addressing this fact for the moment). This is a) *implicit cognitive consciousness*. Or, we can *attend* to this fact. This is b) *explicit cognitive consciousness*.

2. *Phenomenal consciousness* is the directly lived momentary experience. This can be *unattended* (e.g. the sensations of the mouth movements when we are involved in a conversation). This is a) *implicit phenomenal consciousness*. If we *attend* to our mouth movements, we have b) *explicit phenomenal consciousness*.

The word "cognitive" will from here on denote cognitively conscious processes solely. Some scholars use this word for all mental activities, but in this investigation we need to separate the phenomenal level from the cognitive level.

Perception is often described in logical terms as if there were a little man in the brain taking logical decisions. The appraisal theory of emotion is an example of this. This so called intellectualism was criticized by Merleau-Ponty (Merleau-Ponty 1996). There is no proof, he exemplified, that we draw conclusions about distances from lens accommodation, apparent size of objects, from disparity between retinal images, from the convergence of the eyes etc. How could we? This information does not reach the attention of the observer. Merleau-Ponty firmly denied that associations, memories and judgements are involved in perception. But at a closer look it is obvious that Merleau-Ponty treats these faculties in their explicit capacity. His scientific territory was the phenomenally conscious. It was not fully recognized in the 40s that most mental activity is pre-attentive. For example pre-attentive decision taking was shown 1985 by Gazzaniga in his presentation of his split-brain experiments

(Gazzaniga 1985)[50]. According to Gazzaniga almost all decisions are pre-attentive and most logical explanations are rationalizations – logical reconstructions.

The pre-attentive level has been studied with the ERP (event related potentials) technique. This technique registers brain reactions on deviances from expected outcomes. These are pre-attentive reactions. They reveal information about the so-called "primitive intelligence" (Näätänen *et al.* 2001)[51] (for ERP overview, see Tervaniemi, 2003)[52]. The studies show that we perceive, remember and compare (judgement) on a basic level without our knowing. "Primitive intelligence" processes more information than we can handle consciously. And, at a faster rate. These implicit activities are constantly going on under the surface of attention. Even action can take place at an implicit level as when we follow a path while focused on a conversation. In the following perception and perception-action will be treated as *processes* on any level of consciousness.

Can pre-attentive levels create meaning? Yes, in the sense that we implicitly act according to the information. If we, to take a case from Merleau-Ponty's collection of examples, perceive the intervening objects between ourselves and the object attended, these objects will make the distance to the attended object look longer. We act accordingly. We never try to reach for objects perceived to be far away. The intervening objects thus mean distance in the sense that they affect behaviour. There is no calculation on a cognitive conscious level behind this. It is implicit meaning. The distance is perceived. Merleau-Ponty admits that this meaning exists. He calls it immanent meaning "which is not clear over itself".

By the same token there is meaning in music perception. This does not imply that the perception stands for – signifies – meaning. Music can signify, and if so, create explicit cognitively conscious

---

[50] Gazzaniga, M. (1985) *The Social Brain*. Basic Books, New York

[51] Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. (2001) "Primitive intelligence" in the auditory cortex. *Trends in Neurosciences,* Vol. 24. Issue 5. 283–288

[52] Tervaniemi, M. (2003) Musical sound processing: EEG and MEG evidence. *The Cognitive Neuroscience of Music*. (Eds. I. Peretz & R.J. Zatorre) Oxford University Press, New York. 294–309

meaning. But the point here is that there is implicit meaning in the perception of music.

Perception

Perception is about how the world presents itself to us. A. D. Smith states in *The Problem of Perception* (2002)[53] that perception places the object in three-dimensional space. I will argue that this can be generalized to the statement that perspective defines form.

A first step is to show how meaning affects perception. This is an example of so called top down processing. According to Gestalt psychology we perceive relations (Sundqvist 2003)[54]. This is especially true for the perception of music. The general principle in Gestalt psychology is the law of *Prägnanz*. This word means conciseness. "Prägnanz" refers to our tendency to choose the regular and the simple over the irregular and the complex. The most basic relation is the *figure-ground* constellation where the figure stands out from the background as a meaningful coherence. Other Gestalt principles governing perception are the law of closure, the law of similarity, the law of proximity, the law of symmetry, and the law of continuity. The effects can been demonstrated in simple and convincing experiments. Let me take an example from music:

If a melody contains large leaps every second tone, the brain tends to assume that the tones heard in the high register and the tones heard in the low register belong to two separate melodies (Bregman 1990)[55]. This illusory effect makes it possible for one flute to play two voices simultaneously.

We are not aware of the sensory input as such. We can only become aware of what we make of it. As can be seen from the following quotes, perception can be amazingly rich:

---

53 Smith, A.D. (2002) *The Problem of Perception.* Harvard University Press, Cambridge, Massachusetts

54 Sundqvist, F. (2003) *Perceptual Dynamics – theoretical foundations and philosophical implications of Gestalt psychology.* Acta Universitatis Gothoburgensis, Gothenburg

55 Bregman, A.S. (1990) *Auditory Scene Analysis – perceptual organization of sound.* MIT Press, Cambridge, Massachusetts

> One sees the hardness and brittleness of glass, and when, with a tinkling sound, it breaks, this sound is conveyed by the visible glass. (Merleau-Ponty 1996: 229)

> If I stick my finger into a pot of jam, the sticky coldness of this jam is a revelation of its sugary taste to my fingers. (Sartre 1958: 186)[56]

> The myth holds the essence *within* the appearance; the mythical phenomenon is not a representation, but a genuine presence. The daemon of rain is present in each drop which falls after the incantation, as the soul present in each part of the body. Every apparition is in this case an incarnation. (Merleau-Ponty 1996: 290)

The authors apparently mean that these cross modal experiences are not conscious associations but directly felt or experienced in the perception. In the last case even myth is implicated in perception. The presence of the daemon is actually experienced in the raindrop.

The concept of perception has various implications depending on theoretical context. In neuroscientific literature it refers to the processing of sensory input leading to identification of objects; phenomenology treats perception as something momentarily lived; and perception-action theory sees perception as interaction between the world and the active body. This theory encompasses pre-attentive levels of perception.

Several writers have proposed that the purpose of perception is to inform the organism about the *use aspect* of a certain object in a certain situation. Kurt Koffka, one of the fathers of Gestalt psychology, called this the "demand character" of an object: Eat me! Drink me! Fear me! (Koffka 1935: 361)[57].

Close to this concept is James Gibson's *affordance* indicating that the environment affords use-aspects. What we perceive is not the external world as such, but the affordances the external world "makes" us (Gibson, J.J. 1986)[58]. The living organism "picks up" vital information from the environment. The bodily state of the individual

---

[56] Sartre, J.-P. (1958) *Being and Nothingness*. Methuen, London

[57] Koffka, K. (1935) *The Principles of Gestalt Psychology*. Lund Humphries, London

[58] Gibson, J.J. (1986) *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, Hillsdale, New Jersey

(hunger, thirst, fatigue etc.) is of course also important for the recognition of goals or affordances.

Merleau-Ponty used the concept *telos* (Greek for goal) to designate intentional meaning of perception.

All these attempts to understand perception entail that perception is an interaction between the perceiving subject and the environment. The use aspect can be seen in the light of evolution. Every species needs to know how to act to optimize a given situation. The perception apparatus is "designed" to make this possible. Thus different species have a perception apparatus designed to meet species-specific *needs*. The bat has an apparatus for hunting in the dark, fish can detect changes in water pressure, birds perceive directions for flying, flies like the smell of the cow's derriere etc. Every species has a unique filter for olfactory perception that sorts out what is useful from what is not (Buck 2004)[59]. It is the useable aspect of the world we perceive and this aspect is species specific. The bee would not perceive music from musical sounds as little as we would perceive the direction to nectar from the movements of a dancing bee. In sum: The needs steer perception, these needs are species specific and every species has a perception apparatus designed to detect the particular kind of information guiding the animal to satisfy the need.

This is not to say that we perceive the use aspect of every situation. Music or art is not useful in any obvious Darwinian sense. But we perceive it through processes developed to discriminate use aspects.

The situation thus produces meaning in the limited sense that it contains action "proposals". Generally the brain prefers higher order perception: structures are perceived as Gestalts, as affordances, and as *situation understanding*. What then is a situation? I will treat the situation as the total impression of the present moment: sensory input from the surrounding and from the body. This could be called *the sensory field*. The concept field is taken from Merleau-Ponty's field theory categorizing information as:

---

[59] Buck, L. (2004) Nobel Minds, TV-interview. SVT/BBC

1. The sensory field, i.e. the sensory input from the surrounding as well as from the own body.

2. The perceptual field, i.e. all perceptions (processed sensory input) at a certain moment on every level of consciousness.

3. The mental field i.e. memories, Gestalts, internal models.

4. The phenomenal field, i.e. the figure/ground constellation (as in seeing an object in three dimensional space or hearing a melody in a harmonic structure), which constitutes the phenomenally conscious perception. A phenomenal perception is a present now perception.

The situation entails a context – a *frame of reference* – in the perception process. The frame of reference consists of information from the perceptual field and/or the mental field and/or the ground with relevance to the figure. An important frame of reference is the own body. We implicitly "measure" the world with our active bodies. We do not only perceive what is immediately good or dangerous for the body but also instrumentally important features such as weights, directions, speed, and distances. Something is heavy if we have trouble lifting it; it is far away if we cannot reach it; it is fast if we cannot catch it and so on. Merleau-Ponty called this frame of reference *the body schema* (Merleau-Ponty 1996). This includes the implicit feedback of the location of the body parts and the action potential of the body. It involves not only the body but also spatial directions (over/under, left/right, on/under and depth) for intended action. An important factor in the formation of the body schema is gravity, giving a subjective experience of the vertical, without which over/under and left/right would lose meaning. Animals are able to perform accurate movements in their environment. To perform these actions they must have a sophisticated system of commands and feedback involving the own body and the space around it. The idea of the body schema as the frame of reference is supported by experimental evidence: When, for example, a muscle is stimulated with vibrations, the resulting reflexes depend on the position of the body (Berthoz 2000)[60].

The body schema also includes instruments and tools used in bodily action. They become part of the *habitual body.* For example the stick of the blind man functions and is perceived as an extended limb. The instrument of the skilled musician, thus, is included in his body schema. Merleau-Ponty exemplifies this with the organ player. It is impossible to consciously calculate the coordination of hands and feet to bring out music from this highly complicated instrument, he argues. The body-organ connection must be automated. The movements generate bodily and musical feedback. This connects body movements with music. When, in a second step, the organ player adapts to an unfamiliar instrument, this is just like the automatic adjustments we make to talk with food in the mouth, to adapt walking to the terrain, or to drive an unfamiliar car. This proves, according to Merleau-Ponty, that the instrument has become part of the body schema.

The own body can thus be a frame of reference in the perception process. This can easily be shown with reversed goggles, so called because they reverse the up/down or the left/right direction. At first, subjects are confused, but after a while of probing the environment – moving around and touching – they regain normal sight. The brain re-reverses the perception. Another example is provided by Bach-y-Rita. Bach-y-Rita placed a television camera on blindfolded subjects. The picture was transmitted to vibrations on the skin of the subjects. After a while of training how to navigate in the environment, guided by these signals, the subjects learned to "see" with their skin (Bach-y-Rita 1989)[61] in the sense that the signals helped them to navigate. These experiments show that what we perceive is not the actual surrounding but the use aspect of it. We learn to implicitly interpret the sensory input in everyday situations. There is thus no one to one image of the environment in our brains. Not even temporal processes have a one to one representation. ERP-experiments

---

[60] Berthoz, A. (2000) *The Brain's Sense of Movement.* Harvard University Press, Cambridge, Massachusetts

[61] Bach-y-Rita, P., Collins, C.C., Saunders, F., White, R. & Scadden, L. (1969) Visual substitution by tactile image projection. *Nature.* Vol 221. No. 5184. 963–964

indicate that the time is compressed in sound representation (Yabe *et al.* 2004)[62].

In order to use the own body as a frame of reference we need to move. This makes perception an active process. We "scan" the environment with our bodies. What we perceive is not so much information as changes of information. Pressure cannot be felt but the change of pressure is felt. A constant unchanging tone could in the long run not be heard and a picture could not be seen if our gaze is constantly locked in one direction (Boeree 2006)[63]. We need to move to experience change and thus to perceive. As we will discuss, body movement and music perception is an evident example of this.

Another important frame of reference is the memory. We cannot even become aware of music as music without the involvement of memories (Snyder 2000)[64]. However there are different kinds of memories. We have *long-term memories*, *short-term memories* (or *working memory*) and *sensory memories* or *buffers* (Malmgren 2004)[65].

Another categorisation is *procedural memories* referring to the implicit memories of body movements (such as riding a bike), *declarative* (or *semantic*) *memories* (such as facts taught at school) and *episodic memories* – which occur as short reminiscences of lived experiences.

Procedural knowledge is *knowledge-how,* as opposed to declarative knowledge or *knowledge-that* (Ryle 1949)[66]. Merleau-Ponty described a case which shows the difference (Merleau-Ponty 1996):

A man, suffering from brain injury, could not, when asked to, put his hand on his head (although he clearly understood the instruction), but when a fly landed on his head, he smashed it. The conclusion is that although he could not do it deliberately, his body knew

---

[62] Yabe, H., Matsuoka, T., Sato, Y., Hiruma, T., Sutoh, T., Koyama, S., Gunji, A., Kakigi, R. & Kaneko, S. (2005) Time may be compressed in sound representation as replicated in sensory memory. *Neuroreport.* Vol. 16 No 2. 95–98

[63] Boeree, C.G. (2006) http://www.ship.edu/~cgboeree/genpsyperception.html

[64] Snyder, B. (2000) *Music and Memory.* MIT Press, Cambridge Massachusetts

[65] Malmgren, H. (2004) Why the past is sometimes perceived, and not only remembered. *Philosophical Communications, Web Series*

[66] Ryle, G. (1949) *The Concept of Mind.* The University of Chicago Press, Chicago

how to do it, when the lived situation demanded so. Procedural knowledge is the knowledge how to walk, ride a bike etc. – actions we are normally not conscious of. It is the knowledge how to move the mouth to produce words but not the knowledge to understand what is said. It includes knowledge how to reach, catch and find our way in the terrain. Some procedural knowledge, like the newborn calf's ability to walk, is inborn, some is learnt by imitation, and some is consciously learned (declarative) but becomes automated (procedural) after practice. Most of these abilities are too complicated to be taught by communication of signs, or even to be consciously understood. They contain too much information.

These types of memories can be connected with the earlier presented levels of consciousness. Declarative memories are cognitively conscious, episodic memories are phenomenally conscious and procedural memories are unobtainable. Let me exemplify the last statement. We cannot obtain the memory of how to ride a bike. The body knows how to ride, but asked to explain, we do not know how it is done. For this reason, procedural memory is not a memory in the sense that it can be recalled to consciousness. Bicycling is a sequence of movements where every movement is actualized by the preceding movement. And vice versa, even if the bicyclist is a physicist and knows everything there is to know about centripetal force (declarative knowledge), this would not help her on the bike, if she does not have the procedural memory.

Playing music is procedural knowledge. Knowing everything in a declarative sense does not make a good performer. It takes practice. And when we know how to play the piece, it is in our fingers and we do not have to take conscious decisions how to play the next tone. Playing is a movement pattern and once it is established this memory is very robust.

Eleanor Gibson introduced the concept *perceptual learning,* for pre-attentive adaptations to the environment (Gibson, E. 1953)[67]. This knowledge makes us perceive better. Perceptual memories are for the most part learned in early childhood. We develop the ability to perceive from day one, and in fact even earlier. Some of this

---

[67] Gibson, E.J. (1953) Improvements in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin.* No 50. 401–431

learning occurs at certain ages in so called "windows". Eleanor Gibson listed six categories of perceptual learning.

1. *Attention weighting*, which refers to our ability to unconsciously direct our attention to features that are important to us. Example: faces.
2. *Stimulus imprinting*. Through imprinting, detectors (or receptors) are developed to specialize for a certain stimuli. We are much better (faster) at identifying stimuli that we have been exposed to previously.
3. *Feature imprinting* is the ability to select a crucial feature important to identification.
4. *Topological imprinting* is the ability to recognize whole patterns and even to generalize between dimensions.
5. *Differentiation* is much improved by exposure of stimuli. An example of differentiation of complex stimuli is that Caucasians identify Caucasian faces better than Chinese faces.
6. *Unitization*. Parts that co-occur in a set of patterns will tend to be bound together. Evidence of this is our ability to read words as wholes instead of letter by letter, and that faces are immediately recognized without analysis of facial features.

As can be seen from these categories they are relevant to music learning. Let me just briefly comment on feature imprinting and unitization. Feature imprinting is important for our ability to recognize voices and instrumental sounds. Unitization is our ability to link this feature to everything else in the learning situations. A modern term for this is *pattern completion*. This mechanism makes us perceive situations rather than details.

It follows from the fact that perception implies action, that Eleanor Gibson's definition of perceptual learning as a learning how to perceive better, implies the learning *how to act* better to perceive, which to some extent makes this memory procedural. For example the accommodation of the eyes is procedural knowledge which contributes to the perception of depth.

Thus declarative memories and episodic memories are accessible whereas procedural memories are not. This said, it is important to state that all kinds of memories can affect us implicitly. The fact that a memory is accessible does not imply that it is given attention in the perception process. Mostly it is not. If the memory processed in perception is not given attention it is impossible for the listener to separate it from the perception. The memory affects the perception but the listener is not aware of the association as such.

Is this the explanation to the statement by Merleau-Ponty that the daemon can be present in the raindrop? Someone believing in daemons senses the presence of something she "knows" to be there. This can be as vivid as a child's experience of beasts under the bed. Parents know how hard it can be to assure a child that there is no beast in the dark. We light up the room and look under the bed with the child. We explain that the beast is a fantasy. In fact this is to make the implicit meaning explicit – to make the child aware of her own thoughts. It is a way to make phenomenal consciousness cognitive. If we are not aware of the thought as a thought – we have no place to put the beast but in "reality". The beast exists.

Let me now give some examples of how referential meaning affects perception:

1. Last summer I was strolling along the seashore on Öland, an island in the Baltic. From a distance I saw pine trees at the cape. But on closer inspection I realized that it must be masts on a shipwreck. Even the "masts", however, turned out to be an illusion. At closer range I identified these "masts" as poles for fish net drying. What happened here? We do not see the sensation (the input on the retina). The brain builds the image from all information actualized by the situation. Since there were pine-trees all along the shore the perceptual context included these trees. The picture produced by the brain was new pine-trees. Then new contextual information was added – a shipwreck. Now in an instant I saw the masts. This could theoretically go on for ever. The same input can generate different pictures as

the context is changing. As I became aware of the fishing nets I was presented with the image of poles. In this example the frame of reference is the ground (in the figure/ground constellation).

2. Example no 1 is an illusion. It never happened to me. I realized as I reread Merleau-Ponty that the example must have been inspired by him. In this example the frame of reference is the mental field (declarative memory).
3. Seeing may be altered by suggestions by a suggestive person (Raz, Fan & Posner 2005)[68]. This exemplifies that perception can be changed by declarative information.
4. We can see the moon as a flat object. Probably we learned to see the flatness in our childhood. But thinking about the astronomic fact that the moon is a sphere, we can suddenly see it in 3D. In this case the flat moon experience depends on perceptual memories and the round moon experience depends on declarative information.
5. Children see Santa Claus, not a man in a mask. My three year old daughter therefore had an amazing experience when I offered Santa Claus a snifter. He had to remove his mask a little. "Daddy, there was a man *inside* Santa". In this case mythology is the frame of reference.
6. Gestalt psychology uses ambiguous pictures to show that we perceive totally different forms depending on what we take to be figure and what we take to be ground. The same line could be perceived as this or as that, depending on assumed content.
7. The random ink-patches in Rorschach-testing are perceived differently by different subjects. This is comparable to our ability to see animals and faces in clouds. The variation is

[68] Raz, A., Fan, J. & Posner, M.I. (2005) Hypnotic suggestion reduces conflict in the human brain. *Proceedings of the National Academy of the United States of America.* Vol. 102. No 28. 9978–9983

> due to the mental field (which is why Rorschach-testing is used to diagnose mental stages).

Notably the change in perception of form is abrupt. We either see something as this or as that. There are no mediating stages.

Now the reader might ask: If the same drawn line can be perceived as the contour of a profile or the contour of a vase, is this not an example of different content and the same form? No, if we asked someone who perceived the line as a profile to draw it, it would most certainly come out differently than the same line perceived as the contour of a vase. Gestalt experiments show that meaning really changes the perception. The perceptual context can bend straight lines, make equally long lines unequal, change size and add details that are not really there. Or, to take a well known example: the moon is perceived as bigger when it comes close to the horizon. The ground, in this case the horizon, enlarges the moon.

The neurological explanation is that the brain has developed associative connections. We perceive things as this or that depending on which connections are used. These connections can be cross modal. For example lip reading cues to hearing (Pekkola *et al.* 2005)[69].

These are examples of top-down processing. There is a philosophical problem of how we can have a top-down understanding. How can we understand the detail from the situation when we need the details to build up the situation? There are two possible explanations to this question. One is using the concept *internal models* (Klapuri 2001)[70]. Do we have such models? Neurologically this is either inborn so called pre-wired connections or connections that have developed through priming (learning effects). As already discussed unitization (pattern completion) is a pattern of neurological reactions that connects specific features with situational circumstances. Thus if we are to talk about internal models it should be in the sense of potential chain reactions. These patterns connect situation relevant information to an understanding, which includes the

---

[69] Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I., Mottonen, R., Tarkiainen, A. & Sams, M. (2005) Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroimage.* Vol. 16. No 2. 125–128

[70] Klapuri, A. (2001) Top Down Processing. http://www.cs.tut.fi/~klap/iiro/topdown.html

expected outcome of the situation. For this reason top-down processing is sometimes referred to as *prediction-driven* processing. Pattern completion may however not be the only explanation. It has been shown that the perception of a face or an object as well as to the emotional expression of that face are not just assembled from many neuronal reactions putting the details together but also that we have single neuron representations (Fried *et al.* 1997)[71]. This indicates the possibility that we might have "situation neurons" (single neuron that can detect a typical situation).

Either way this is implicit situation understanding. In the case of music it would be the sense of a certain style.

To sum up:

1. The situation is the input of internal and external information. This is the sensory field.

2. The perceptual field consists of all perceptions on any level of consciousness.

3. Situation understanding is basically expectations (derived from the frame of reference) and reaction tendencies. This is the implicit meaning of the situation.

4. The figure attended to (or form) is a "proposal" by the brain of what we are facing based on situation understanding (content). This figure belongs to the phenomenal field.

The frame of reference thus provides us with information which affects perception of the perceived form. The point I want to arrive to is this: When the meaning is implicit, it cannot be separated from the phenomenally perceived object.

---

[71] Fried, I., MacDonald, K.A. & Wilson, C.A. (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron*. Vol. 18. No 5. 753–765

Application to music

The presented theory entails that music is perceived differently between listeners. Depending on personal background, exposure (perceptual learning), mood etc. some people hear things that others do not. It is sometimes said that we all perceive music the same way. But this is because all westerners are exposed to enormous amounts of Western popular music each day. If we were to listen to Vietnamese classical music, would we perceive it the same way as a performer of that music? The listener rarely exposed to Vietnamese music has no internal model of it, no sense of the style, and thus no expectancies. To her the music has no meaning. This deteriorates the perception. She would recognize it as music because of some universal features, but if asked to repeat the melody, she would face a problem.

Since most studies of perception treat visual perception, special attention must be paid to the question if the findings are valid for audible perception. There are some fundamental differences between seeing and hearing. We can, to take a concrete example, hear but not see around corners. Thus the perception of space and perspectives is fundamentally different in hearing. We do, however, in other respects receive information about spatiality, directions and movements from audible cues.

Hearing is more temporal than seeing. Sounds generally have shorter duration than images. Everything we do produces sounds. This is important to our sense of timing. Audible cues, such as the sound of the tennis ball, are important to our ability to time our own body movement to get a clean hit. Even diminutive sounds as those produced as I type on my computer, are important to my sense of touch with reality. Out of sync experiences are very disturbing (as the delay in speech in intercontinental phone calls).

The concept timing refers to how events are spaced in time. In a second step we can think of timing as the adjustment of body movements to environmental events and environmental structure (as the timing of a downhill skier is adjusted to the gates). Obviously there is timing between the body and the music structure. Let me present some arguments that this is the case:

1. Music makes us move. One of the main functions of music is to produce and support body movement. We can see this from work songs to gym work out, from ancient rituals to break dance.

2. Listeners move, not only to the rhythm but sometimes even to the melodic "gestures". Although there is no obvious direction of melody movement, most listeners feel directions in music. When the melody is moving "upwards" or "downwards" you get a feeling of spatial direction.

3. We can perceive the contour of the gesture depending on how it is played: legato, staccato - as smooth or uneven.

4. We can perceive the distance between the notes: "small steps" or "giant steps".

5. The music/movement connection is used in film music (Bryngelsson 2006)[72].

6. We use, as can be seen above, body metaphors in our description of music. This indicates that we can perceive music with the body schema as a frame of reference.

7. Most music, even classical, is derived from dance styles. There is a reciprocal relation between music and dancing. Not only are dance movements adapted to music but the music is adapted to dance movements. Thus music is adapted to the human body – to its moving potential. We sense this movement pattern when we hear a gavotte or a samba, even if we are just vaguely familiar with these dance styles.

Why does music make us move? Music listening and even imagining engages pre-motor centres in the brain. Researchers of movement such as Berthoz (2000) take an interest in synergy effects – the coupling of movement commands to perceived sequences.

---

[72] Bryngelsson, P. (2006) *Filmmusik – det komponerade miraklet*. Liber, Malmö

The other side of this coin is that music itself is movement. A melody moves upwards or downwards, moves in steps or as a gesture. Do we perceive music visually – as a spatial structure? Some people do. This question will be treated in Chapter 4.

Are the demand/telos/affordance aspects applicable to music perception? Can we talk about a situation understanding? At first sight we do not seem to have a strong case. Music, as well as other aesthetic art forms, does not have any obvious survival value. It is not in any obvious sense demanding anything from us; it does not afford any obvious use or present us to a situation with any obvious goal understanding. This problem is reflected in the variety of developmental explanations of why we have music. If music is inviting us to anything, it is an "Aufforderung zum Tanz" (invitation to dance), or more generally an invitation to move. But what is the developmental explanation?

There is a possible answer to this question. As stated in the introduction, music can been considered a target domain – a composite of source domains. These source domains have survival value. This simply means that we enjoy music with a perception apparatus "designed" for basic use purposes.

Form & content

As stated in the introduction, the formalism contra representation (of content) debate has been a watershed in the philosophy of music. In order to connect this thesis to the debate, I will make an effort to discuss perception and meaning in form/content terms. How do form and representation relate to perception and meaning?

An influential understanding of form and content was presented by the Danish linguist Louis Hjelmslev in 1943 (Hjelmslev 1993). Linguists separate the semiotic side from prosody in language. While semiotics in language is the symbolic relation between the words and their denotation, prosody is principally the melodic quality of the sentence – the expressive delivery. What is form and what is content here? Word sequences are forms and so is prosody. While words represent the content in the sense of what the speaker wants to communicate, the prosodic expression carries information about identity, accent and background of the speaker. And what is more, prosodic expression reveals emotion, intention, style and personality. We even judge if the information is trustworthy by the way it is delivered. So, both word sequences and prosody carry information.

Hjelmslev must have thought of this when he stated that content and expression generate two kinds of form in speech – *form of content* and *form of expression*. But, and this is the point, since a sentence objectively can have only one form, Hjelmslev must have referred to form as the form subjectively perceived. Otherwise it is hard to defend that there would be two forms.

If we take form to mean the form as perceived by a subject, Hjelmslev's proposal can be interpreted according to what has previously been said about perception and meaning. If we try to understand what someone is telling us, the word sequence will be attended to, while the prosody remains unattended to. But, if we choose to focus our attention on the information carried by the expression, the prosody will be attended to while the sequence of the words will remain unattended to. The figure and ground will be reversed. Just as in the Gestalt pictures this is due to the information implied. Thus two different forms can be derived. Since there are two kinds of content here, Hjelmslev's categorization is somewhat unsatisfactory. For this reason I would specify that *form of content* is *form of semiotic content* and *form of expression* is *form of expressive content*. I propose that this can be generalized to the statement that *form is content dependent*. This means that we do not only judge the content from the form, but also that we implicitly judge the form from the content. Form and content presuppose each other.

A sign is something, which stands for something to somebody (Peirce 1931: Vol. 2 § 228)[73]. As stated before the meaning can be implicit or explicit in the perception. As I have already stated: If the information is implicit in the form, form and content cannot be separated. The form cannot stand for something to the observer. Under these premises there is no representation – the content is perceived

[73] Peirce, C.S. (1931–58) *Collected Papers*. (Eds. C. Hartshorn, P. Weiss & A.W. Burks). Harvard University Press. Cambridge, Mass

as form. I intend to show that this is largely the case with music. This is, to talk with Peter Kivy, enhanced formalism.

My understanding of *form* in music is *a recognizable musical structure*, since some sort of recognition is needed to perceive it. And the *content* in music is *the implicit or explicit understanding that makes the musical structure recognizable*.

According to this theory the perception of music is not in everyday listening an interpretative process where we draw conclusions about the content from the form.

Imagery

Imagery traditionally refers to the phenomenon that the brain can produce images or sounds without external input. It is a "quasi-perceptual experience", meaning that it resembles perception. Most people have this experience. As we take a walk, a melody can present itself. It rings in your head. How does this work?

Thomas (1999)[74] offers an intriguing theory. Illusion, Thomas argues, can be said to be the same thing as *imagination*. Imagination is a wide term with a long history in the literature of philosophy. It has become a somewhat vague concept during this journey.

Imagination can be triggered by events in our surrounding. We imagine what we see. This means that we take it to be this or that. We imagine consequences, and we imagine the intention of the other. It is a situation understanding.

Imagination can also be triggered by memories. We stage situations and imagine: *What if* this or that would happen. It is a play with possibilities and possible outcomes. Imagination thus prepares us for eventualities.

Thomas sees imagination as the production of imagery. His understanding of imagery entails a continuum from *seeing as* to *clear-cut imagery*. *Seeing as* is an implicit interpretation of the input whereas

---

[74] Thomas, N. (1999) Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*. Vol. 23. No 2. 207–245

clear-cut imagery is a "perception" right out of the blue. The idea of a continuum makes imagery an agent in the perception process.

It has been experimentally shown that imagery actually intervenes with perceptive organs. Efferent (outgoing) fibres from the brain to the retina provoke retinal reactions (Repérant *et al.* 1989)[75]. The efferent fibres comprise about 8% of the optic nerve. Likewise efferent fibres *from* the brain *to* the ear make the hair cells move (Dallos 1992)[76]. This indicates that we actually experience imagery with our sensory organs. Predictions guide the attention and affect the sensitivity of the auditory system (Klapuri 2001).

Imagined objects thus may be quite real. In fact we handle imagery much the same way we handle real objects. Thus an inner image of a cube can be rotated (Shepard & Metzler 1971)[77]. An inner image of a page can be scanned across (Kosslyn 1973)[78]. Inspecting mental images, subjects find large details quicker than small ones (Kosslyn 1975)[79]. The time required for performing imagined acts is correlated to the time required for performing these acts in vivo. Marc Jeannerod points out that the brain areas activated by perception and imagery overlap:

> [m]ental images, at least in the visual modality, rely on the same neural substrate as the perceptual images that are generated during normal perception. …motor imagery pertains to motor physiology in the same way as visual imagery pertains to visual physiology… motor images must also be considered to be functionally related to the imagined movement. (Jeannerod 1994)[80]

---

[75] Repérant, J., Miceli, D., Vesselkin, N.P. & Molotchnikoff, S. (1989) The centrifugal visual system of vertebrates: A century old research reviewed. *International Review of Cytology.* Vol. 118. 115–171

[76] Dallos, P. (1992) The active cochlea. *Journal of Neuroscience.* Vol. 12. No12. 4575–4585

[77] Sheppard, R.N. & Metzler, J. (1971) Mental rotation of three-dimensional objects. *Science.* Vol. 171, No. 972. 701–703

[78] Kosslyn, S.M. (1973) Scanning visual images: Some structural implications. *Perception and Psychophysics.* Vol. 14. No 1. 90–94

[79] Kosslyn, S.M. (1975) Information representation in visual images. *Cognitive Psychology.* No 7. 341–370

[80] Jeannerod, M. (1994) The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences.* Vol. 17. No 2. 187–245

A convincing example of the connection between imagery and motor functions can be fetched from the field of sport: Yue and Cole (1992) compared the increase in muscular strength produced by actual training (repeated maximal isometric contractions) and by mental training (imagining the same effortful contractions). In both conditions, the maximal force produced by the trained muscles increased significantly (by 30% and 22%, respectively).

The perception-action theorist Ulrich Neisser has clamed that perception can be thought of as a cyclical process (Neisser 1976)[81]. It has been suggested that this process is a "hypothesis fitting" of "proposals" made by the brain of what we are actually seeing (Marcel 1983)[82]. *Seeing* something *as* something then becomes following such a proposal. The idea is that the brain provides the subject with an idea – a hypothesis – and completes the lack of incoming information with its own information about this hypothetical thing. This theory can no doubt be criticized for intellectualism. Of course the brain does not in any conscious way test hypotheses or makes proposals. I propose another metaphor that describes this better.

Let us say that the police have secured a fingerprint at a crime scene and that this print is incomplete. The police will run the sample against a data bank of fingerprints. The computer comes up with a possible match. The screen shows the matching fingerprint and adds relevant data, such as the identity of the person and her criminal record. In this allegory the computer is the brain, the sample is the environmental input, the data bank is the memory, the program is the implicit processing and the screen is the consciousness. The information on the screen is the imagery. Contrary to the input and the program this information is user-friendly. This and only this is what actually will be experienced. Nothing else in this process is experienced. It is in this sense we can talk about a "proposal" or a "hypothesis". And it is in this sense we can talk about imagery as a prediction. The fact that imagery, when more information is available, can be proven right or wrong shows that we are dealing with

[81] Neisser, U. (1976) *Cognition and Reality*. Freeman, San Francisco

[82] Marcel, A.J. (1983) Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*. Vol. 15. 238–300

implicit expectancies and predictions. Since we are not aware of the sensory input, there is no way to notice that the observation partly is made up of information that is not actually on the scene.

Imagery thus presents us to an intelligible user-friendly world. This way of looking at imagery is somewhat contradictory to the definition that it is an image (sound) produced by the brain without external input. But it is consistent with Thomas' understanding of imagery as a continuum from clear-cut imagery to *seeing as*. If there is no external input the imagery usually gets weak but if there is an external input it becomes a strong imagination of what it is we are facing. Imagery thus fills out lacking information. This is called *complementary imagination*. Let me give a few examples!

1. When we listen to speech we do not pick up every syllable, yet these syllables are heard. The phenomenon is called *auditory restoration* (Warren 1970)[83].

2. Visual restoration can also be exemplified. In the eye we have a blind spot, where the optic nerve leaves the eye, but we are not aware of this blindness; the brain fills in predicted content. This is called *illusory continuity*.

3. When a continuous sound is broken, the break, if short, will not be heard (Klapuri 2001).

4. The missing fundamental. A tone from an instrument consists of the tone played (the fundamental Fo) and the overtones, which gives the tone its timbre. Even if the fundamental is masked, listeners actually hear it. The explanation is that we are able to draw an implicit conclusion about the fundamental from the overtone series.

The relevance of imagery to form now becomes obvious. The brain approaches the archive of stored information. Stored information (memories) can be retrieved by neuronal pathways from sensory input to association cortex. These pathways develop through *priming*

[83] Warren R.M. (1970) Perceptual restoration of missing speech sounds. *Science, 23*, Vol. 167. No. 3917. 392–393

(perceptual learning). Priming causes change in the firing potential of the neurons.

If we have sensory input not fitting the suggestion, the librarian runs back to consult another archive. The chosen information is the "hypothesis" of what we hear. A form will appear. This means that Hjelmslev was right. Form in an absolutely concrete sense is made up by content. *Imagery is content perceived as form.*

The difficulty with the concept form is the same as the with the concept internal model. It is too static. We have to include the time factor. This is exactly what Peter Kivy did when he suggested that enhanced formalism implies a temporal topography (Kivy 1989) – music as a moving or unfolding form. The brain has learnt from various situations that if this happens (cue) one of the allowed alternatives will follow (rule). As we move forward in a real time perception the felt continuance becomes a new cue for the next rule governed continuance. Thus imagery unfolds like falling dominos. For this reason the concept *sensory-motor schema* (Deleuze 1989)[84] may serve as the neurological counterpart to the temporal topography.

Music imagery

Even if imagery is involved in all perception and can be found for all modalities, music imagery is probably the most evident kind. It is a well known phenomenon experienced not only by anybody involved in music, but by most listeners that we "hear" the expected outcome in music. Yet it is a new area of research. Pioneering work by Robert Zatorre and Andrea Halpern in the mid 90s has been followed by others.

The most obvious example of music imagery occurs when music is suddenly muted. We will still "hear" the expected continuation. A variant of this is that we feel the beginning of the next song on a CD as the foregoing song is ending.

Some people can hear melodies in their heads even if there is no external source. Here are some examples of music imagery:

---

[84] Deleuze, G. (1989) *The Time Image*. University of Minnesota Press, Minneapolis

1. A musician can rehearse a concert in his mind, as Horowitz and Rubinstein are reported to have done (Meister *et al.* 2003)[85]. This shows that it is possible to sequence and cue a whole concert to a real time inner performed and heard experience. The musician mentally moves through the piece, projecting herself into the music structure. Just as we can imagine following a geographical path, we can imagine following a piece of music sequence by sequence. Meistner found in an fMRI study that such rehearsal engaged just about the same areas as the actual playing, with the exception of primary motor cortex (which could be expected since no actual movements are performed). Premotor areas however are activated.

2. The type of imagery described in the Horowitz and Rubinstein cases concerns a fixed material. But in my experience there is also another variant of imagery, a creative variant that makes people "hear" new melodies. The melodies heard are predictions and these predictions are cues to new predictions. This way, whole pieces of new music present themselves. Unknowingly I have always used this in my own composing. I work in real time. I follow the "proposals" from my inner hearing. When I come to a stop, I go back to the beginning and, so to say, start the mental tape again: "replay". When I reach the stop, the continuance usually presents itself. Musicians asked to explain how they compose often say: "It just comes to me". Sometimes they refer to God or to some kind of antenna. A composer may perceive herself as a medium. They just pick up what is presented to them. Composing is to a large extent intuitive.

3. We tend to group sounds into musical structures (Gestalts). When we listen to the ticking of a clock, we might start imagining a melody. If we hear random beats, our brains

[85] Meister, I.G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R. & Thron, A. (2004) Playing piano in the mind—an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research*. Vol. 19. No 3. 219–228

suggest rhythmic groupings. Thus non melodic sounds and random sounds can be perceived as music because the brain "suggests" it is.

4. There is a strong transfer effect between lyrics and melody (Ullén *et al.* 2005)[86]. Words cue for melody and melody cues for words. You can test this by reading Happy Birthday rhythmically and see if you "hear" the melody. Then sing the melody and see if you "hear" the lyrics. It is impossible to avoid this implicitly learned connection.

5. Rehearsing the finger movements on a table, triggers music imagery (Repp 2001)[87].

6. Pietro Rizzo, Chief Conductor for the Gothenburg Opera says in an interview that when he reads the score, he hears the music in his head. The violins, the timpani, the trumpets – every instrument and all voices sound exactly as he wants them to – "…then, as I approach the orchestra, I listen to their playing and compare that to what I hear in my head. I register where it differs and try to change that" (Rizzo 2006)[88].

As we have seen from these examples music imagery can be clear-cut imagination as when we "hear" music in our heads without an external source, it can be elicited by sounding music, by lyrics, movements, and in the case of experienced musicians it can be triggered by music scores.

---

[86] Ullén, F., Bengtsson, S., Ehrsson, H.H. & Forsberg, H. (2005) Neural control of rhythmic sequences. *Annals of the New York Academy of Sciences.* Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo)Vol. 1060. 368–376

[87] Repp, B.H. (2001) Effects of music perception and imagery on sensorimotor synchronization with complex timing patterns. *The Biological Foundations of Music.* (Eds. R.J. Zatorre & I. Peretz) *Annals of the New York Academy of Sciences.* Vol. 930, June. 409–411

[88] Rizzo, P. (2006) Göteborgs-Posten, October 14

Imagery is the only way we are able to grasp structure in music (Halpern 2001)[89]. We cannot embrace the whole structure in one thought, or jump directly from one place in the music to another. We can only imagine it as played – in the same time direction and in approximate real time.

Music imagery implies activation not only of auditory centres but of motor centres such as the supplementary motor area (SMA) involved in motor planning (Halpern 2001) and the sensorimotor cortex (Schürman *et al.* 2002)[90]. This indicates that we not only imagine the sound of music but to some extent even the actual performing.

Schürman, Raij, Fujiki & Hari (2001) provoked music imagery by having trained musicians read visual notes. Magnetoencephalography revealed that first occipital areas were activated and then the auditory association cortex. Finally the sensorimotor cortex combined with the inferotemporal association cortex. A possible interpretation is that note reading leads to a complex temporal imagery including hearing these notes sound, followed (within milliseconds) by latent motor playing impulses and visual associations to the sight of the keys of the instrument.

The observation that pre-motor centres are involved opens up for an interpretation of music imagery. Consider the following case: A subject whose oculo-motor muscles are paralysed sees objects moving to his left whenever she believes she is turning her eyes to the left (Merleau-Ponty 1996). This qualifies as imagery since she experiences something that is not grounded in sensory input. My interpretation is that the pre-motor centre is activated by her intention to look to the left. The motor schema causes the imagery and makes her believe that she is moving her eyes.

Can this be applied to music perception? Can imagery emanate from movement schemas triggered by actual or recalled music? As the music proceeds, the body takes part actively or latently. I am referring exactly to those pre-motor commands noted at brain scans

---

[89] Halpern, A.R. (2001) Cerebral substrates of musical imagery. *The Biological Foundations of Music.* (Eds. R.J. Zatorre & I. Peretz) *The Annals of the New York Academy of Science.* Vol. 930. 179–192

[90] Schurmann, M., Raij, T., Fujiki, N. & Hari, R. (2002) Mind's ear in a musician: where and when in the Brain. *NeuroImage.* Vol. 16. No 2. 434–440

of music perception as well as of music imagery. This pre-motor activity can theoretically produce expectations that are perceived as actual sounds. If so, imagery is a procedural phenomenon. We would not have to remember how the music develops; we just have to follow the movement schema. Less strongly formulated I suggest that there is a reciprocal relation between imagery and pre-motor activity. This would explain why we can only remember music in real time sequences. This argumentation supports the suggestion that music memory is a sensory-motor schema.

The connection to movement suggests that we latently imitate the movement of the music and that this makes us perceive form. Familiarity to the music might be a condition for our ability to imitate, since familiarity creates expectations (of the next tone/move). Familiarity demands perceptual knowledge and internal models. This is implicit content. This indicates that the link from content to perceived form in music is movement. This theme will be developed further.

Perspectives

I have argued that the same form can be perceived differently depending on the content (meaning). How is this information selected?

> There is an extraordinary amount of interconnectivity between the thalamus and the cortex that allows the cortex, particularly the forebrain, to regulate the flow of information to itself…, so that it can focus on what it needs to know about from moment to moment. (Sakharov, Davydov & Pavlygina 2005)[91]

This quote expresses that the nervous system makes a situation-adjusted selection of relevant information. This flow constitutes the frame of reference. The body seems to “strategically choose” a frame of reference depending on the task (Bertholz 2000). The frame of reference implies a *perspective*. In the concrete understanding a perspective is taken from a specified *landmark*. Depending on the

[91] Sakharov, D.S., Davydov, V.I. & Pavlygina, R.A. (2005) Intercentral relations of the human EEG during listening to music. *Human Physiology*. Vol. 31. No 4. 392–397

landmark, some information becomes relevant and some becomes irrelevant. This process creates a situation understanding. Every perception thus depends on the perspective.

The choice of perspective is mostly an implicit act, although it can be voluntary. For example we usually see the moon as a flat object, but we can purposely choose to perceive it in three dimensions. We do that by thinking of the fact that the moon is a sphere. Thus we actualize a frame of reference by will. But once we do that the flatness totally disappears. We are just presented with one proposal at a time. The frame generates the proposal (imagery) of what we are facing. This explains the sudden shifts in perception.

The own body is the landmark that constitutes *the egocentric perspective.* From this viewpoint we perceive objects in an egocentric space. The body schema is the frame of reference. Merleau-Ponty (1996) pointed out that the perception of distance is perspective dependent. From the perspective of the observer's own body distance can be length, breadth and height. We have an implicit sense of the vertical, the horizontal and depth as separate entities. We use different muscles for eye movements and head movements to estimate height, breadth and depth. From the egocentric landmark, these distances are not analogous dimensions. Thus the egocentric perspective creates a three dimensional room – the ground in which objects may be perceived.

Allocentric perspectives (lat. lóco: place, Eng: allocate: place out) means as the name indicates that the landmark is allocated to a point in the environment. The allocentric perspective maps the terrain, since it relates points in the terrain to each other. This map becomes the frame of reference.

Landmarks located in other persons constitute *relative perspectives.* The *relative* perspective implies the subject's ability to project into someone else's position. The subject imagines a situation from someone else's landmark. The discovery of mirror neurons is relevant for the understanding of this perspective. The suggested mechanism is identification through implicit bodily imitation. We experience the other through our own bodies. We will come back to this.

I intend to examine music perception from these perspectives. However I will divide the relative perspective into three parts: the *dyadic* perspective implies a frame of reference consisting of the information space shared by two individuals, the *tribal* perspective implies a frame of reference shared by an interacting group, and the *cultural* perspective implies a frame of reference shared by a culture. In the last case the knowledge is coded.

The tribal perspective is constituted by the interaction of the group. This creates a shared perception and thus shared experiences and a shared phenomenal field that is quite distinct from those of the individual perspectives. As the group interacts it co-perceives itself as an interactive identity. This is the tribe. When this group becomes big enough it can no longer interact but it can still communicate with cultural codes. I will discuss these two perspectives in the chapters called *Tribal Rites* and *Contemplating Music*.

When we take a perspective we integrate it with the egocentric perspective. This means even if I take an allocentric perspective, e.g. the landscape as seen from a tree, I have to imagine myself standing at the tree. In the dyadic perspective, I imagine myself in the others position. In the case of the tribal perspective I perceive myself as an interacting member of something bigger than myself namely the group. To take the cultural perspective is to relate events to cultural (declarative) knowledge.

This is a rough classification. In fact each of these five perspectives entails a range of possibilities. For example any other person can be the other in the dyadic perspective; every choice of viewpoint in the terrain generates a specific allocentric perspective; any group in which we interact and any culture in a multicultural society provides specific perspectives. The perspectives listed are strictly speaking rather classes of perspectives than perspectives. However, I will not burden the text with this reservation unless necessary in the following presentation. Thus we have the following definitions:

1. The egocentric perspective: the landmark is the listener's body and the frame of reference is the body schema and the perceptual ground.

2. The allocentric perspective: the landmark is a point in the terrain. The frame of reference is the mapped structure.

3. The dyadic perspective: the landmark is the other and the frame of reference is the information space shared by two individuals.
4. The tribal perspective: the frame of reference is the information space shared by a group.
5. The cultural perspective: the frame of reference is culturally coded knowledge.

Needless to say the mechanism of taking perspectives is economizing. We do not have to address everything simultaneously. But also, and this is even more important, it structures our perception and makes it more useable, since a perspective is a selection of information for the situation at hand. We can see how different perspectives make the world accessible. Let me just hint at some implications: The egocentric perspective entails what I perceive from my body. The allocentric perspective maps the environment. The dyadic perspective makes it possible to identify with the other. It opens up the possibility of spontaneous emotional connecting. The tribal perspective takes in the extended family. It is created in interaction with the group. And the cultural perspective depends on cognitive tools to communicate thoughts and knowledge.

As we do these operations we see that the concepts "landmark" and "ground" are replaced. "Landmark" is replaced by "point of view" in a metaphorical sense and "ground" is replaced by "implying this contextual understanding". If we were to define the word perspective it would be: as perceived in a frame of reference. The frame of reference may be the phenomenal ground but it could also be the mental field (memories and internal models) and the state of the body or combinations thereof.

The chosen perspectives reflect that we can understand something with reference to our own bodily, to shared knowledge (as the shared experience of married couple), to the surrounding, to tribal knowledge (as the shared world of a football team) or to our culture with a common language and a common media.

In the following I will discuss the implications of perspective for imagery, form & content, perception, figure & ground, time, space and finally – music.

Perspective & imagery

Experimental evidence has shown that recognition of objects placed on a table is better if the subject has the opportunity to move around the table compared to if the table is rotated (Simons & Wang 1998)[92]; (Wang & Simons 1999)[93]. This has been taken as evidence of a motor program, updating place from the active motion of the subject. But, the same results were achieved if the experiment was performed in virtual reality (Bülthoff & Christou 1999)[94]. The subjects were now shown film sequences from a subjective camera. In a following experiment the subjects just had to imagine that they were moving. Compared with the rotating table both subjective camera and pure imagination showed to be quicker and more accurate for recognition of the objects on the table. In this case there is no obvious motor component (Wraga, Creem & Proffitt 2000)[95].

I believe that this has to be understood in light of the fact that clear-cut imagery is neurologically close to perception. As mentioned before there are indications that pre-motor areas are active during imagination. If we see a film taken by subjective camera or just imagine that we move around the table, pre-motor activity ought to follow. This pre-motor program is activated as we recall the situation. The experiments imply that imagination is the integration of long-term memory of allocentric perspective (mapping of the environment) to egocentric perspective. It is *as if* I was moving in a remembered landscape, perceiving that landscape from my own body.

---

[92] Simons, D.J. & Wang, R.F. (1998) Perceiving real-world viewpoint changes. *Psychological Science.* Vol. 9. No 4. 315–320

[93] Wang, R.F. & Simons, D.J. (1999). Active and passive scene recognition across views. *Cognition,* Vol. 70. No 2. 191–210

[94] Bülthoff, H.H. & Christou, C.G. (2000) The perception of spatial layout in a virtual world biologically motivated computer vision: *First IEEE International Workshop,* BMCV Seoul, Korea, May, Proceedings, Vol. 811. Springer Berlin / Heidelberg. 10–19

[95] Wraga, M., Creem, S.H. & Proffitt, D.R. (2000) Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology. Learning, Memory and Cognition.* Vol. 26. No 1. 151–68

It has been shown that hippocampal neurons indicating place are activated by imagined visual stimuli (Kreiman, Koch, & Fried, 2000)[96]. This supports the idea of allocentric integration in egocentric perspective. It is like moving in a dream in the sense that the dream consists of memories that are always perceived from an egocentric vantage point. In an overview of recent research, Neil Burgess (2006)[97] concludes that "egocentric and allocentric representations exist in parallel, and combine to support behaviour according to the task". Burgess speculates that the egocentric-allocentric division follows the neurological distinction between a dorsal and a ventral processing stream. Burgess then states that "the egocentric nature of perception and imagery" (meaning that we always perceive everything and imagine everything from our own location) requires that input to and output from allocentric systems are mediated by transient egocentric representations.

It should be noted that corresponding dorsal and ventral streams have been identified for hearing. For this reason I propose that Burgess' theory can be generalized to hearing and furthermore that it can be generalized to all perspectives. Imagination thus is to integrate any frame of reference with the egocentric frame of reference. To put it simply: I live this information (perspective representations) *as if* I was there. This fits with Thomas' theory. According to Thomas, imagination is an *as if* understanding. What I perceive in this frame of reference is imagery, because the situation is imagined.

An example: Albert Einstein exercised the imagination what it would be like to travel on a beam of light. This means that he integrated knowledge from physics (cultural perspective) with his *egocentric* (as seen from his own body) perspective. This led him to the theory of relativity, which considers the very fact that the world is perceived differently from different perspectives. It is relative.

This calls for a specification of our understanding of perspective taking. *To take a perspective is to integrate another perspective with the egocentric perspective*. To take a dyadic perspective thus is to see the

---

[96] Kreiman, G., Koch, C., & Fried, I. (2000). Imagery neurons in the human brain. *Nature*. Vol. 408. No 6810. 357–361

[97] Burgess, N. (2006) Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Science*. Vol 10. No 12. 551–557

world as if I were in your shoes. But not quite, since I am aware of my own position and since I am able to separate the inner events from environmental events. I do not lose track of my original position. Taking another perspective is therefore a co-perception of two perspectives – my own and that of the other's.

There are different levels of imagery. Sometimes imagery completes the perception with details that can be proved nonexistent and sometimes imagery is so vivid that it is confused with reality (hallucination). Thomas suggests a continuum. I suggest that we take one step further: Imagery is what we consciously perceive. Always. Perception, imagery and hallucination are just different degrees on a continuum of correspondence with environmental events. The information that appears on the screen of consciousness is the product of processed sensory input and implicit information selected by the perspective. I suggest that this is the function of imagery and the normal case. Clear-cut imagery is an astonishing side effect. In this case only, the inner world (clear-cut imagery) can be separated from the perception of the environmental world. Here we have two perspectives in conflict. Somehow we seem to recognize that. This is evident when we carry out an imagined conversation – an inner dialogue. If I lose track of my own landmark – the other becomes a real presence. This is a hallucination.

Perspective, form & content

The form of an island is different from different ships. And then again, the form is objectively constant. Objectivity expressed in perspective terms is *intersubjectivity* indicating the sum of several egocentric perspectives. If the information from the ships is integrated, a constant form of the island could be assembled. The same constant form would appear if one navigator moves around the island gathering information from different viewpoints. According to perception-action theory this activity of moving around as well as handling the object – picking it up and using it – is how we learn to perceive. As we view something, we implicitly know what it looks like from another direction. We have a sense of the hidden side of a cup, even if we cannot see it from where we happen to be. We feel that the hidden side is rounded. We can imagine this because we have picked up the cup and used it in the past.

A ring can be seen both as an oval and as a circle. In fact we have a difficulty to see it is as an oval even if the retinal input is oval. If someone asks about the form, very few would say that it is oval. But consider this: The ring is oval from an infinite amount of perspectives and round only from two. If somebody asks about the size of the ring, you would not say that it is smaller when it is seen from a distance. This is because the ring is perceived in a three dimensional space. The brain implicitly compensates for the smaller sensory input. The use aspect is also relevant for the perception of form and size. In the case of the ring, we have memories of putting a ring on the finger. This is the use aspect afforded by the object. It cannot very well be oval if it is to fit on the finger, can it? Neither can it suddenly be small. In short: Probing expands the frame of reference and contributes to the constancy of the perception of the ring. In James Gibson's terms the ring becomes *invariant.*

In some respects we can perceive from the very start. Newborns can imitate tongue and mouth movements (Meltzoff & Moore 1997)[98]. They also spend more time watching pictures of faces than abstract pictures (Gelman & Opfer 2002)[99]. This indicates that babies actually have some kind of meaningful perception. Thus memory does not seem to be implicated in all perception. We are obviously pre-wired with a program for perception and then we develop this program for the rest of our lives. There are in fact a number of indications (in relation to the mother for example) that indicate that babies react to situations. Some of these fade away as the child grows older (such as the baby's ability to keep the breath under water).

---

98 Meltzoff, A.N., & Moore, M.K. (1997). Explaining facial imitation: theoretical model. *Early Development and Parenting*. Vol. 6. 179–192

99 Gelman, S. & Opfer J.E. (2002). Development of the Animate-Inanimate Distinction. *Blackwell Handbook of Cognitive Development* (Ed. U. Goswami). Blackwell Publishing. Oxford, UK.151–166

Perspective & perception

Merleau-Ponty was faithful to a subjective experience in his understanding of perception: This is how I perceive it from here. He treated perception as a here and now lived experience that could not be broken down in parts whereas his so called intellectualist opponents confused this direct experience with scientific declarative knowledge in their discussion of perception. Merleau-Ponty argued that all knowledge of the world originally emanates from this exact moment of direct perception. In a way, both sides were right. It takes a perspective theory to see this. Merleau-Ponty's was right in his analysis of perception if we consider the conscious side of perception. But if we consider the preconscious side, perception is implicitly coloured by phenomenal as well as declarative memories. Merleau-Ponty and the intellectualists treated perception from two different perspectives and both sides failed to see it from the other's point of view.

My point is that perception is perspective dependent. As mentioned early in this chapter, A. D. Smith has argued that *perception places the object in a three dimensional room*. This connects to the statement made above that *perspective defines form*. But since the frame of reference encompasses not only the ground but also the mental field and the state of the body Smith's conclusion could be seen as a special case, where the frame of reference is limited to spatial dimensions.

The proposition that perception is perspective dependent is my motivation for a perspective theory of music perception. It is, I suggest, the entrance to enhanced formalism.

Figure and ground in perspectives

As noted earlier in this chapter, imagery is a "proposal" by the brain, a "hypothesis" of what we are facing. The figure presented to consciousness is the winning hypothesis: I see this.

To shift the attention from the figure to something in the background is to shift perspective. The figure is attended whereas the ground is unattended and becomes frame of reference. The line outlining a figure in a drawing is in fact the line separating attended information from unattended information. This coincides with invariance (constancy) contra variance – the perception of a horse, say, against a moving background, if we follow the horse with our gaze.

The figure is understood in relation to the ground. As already pointed out, figure and ground make up the phenomenal field. In this case the ground is the frame of reference. It makes us see the object as something situated in a three dimensional space. But since the three dimensional room is mental (the implicit understanding of the space as three dimensional) and since this mental representation is based on the body schema, the ground is just one aspect of the "frame of reference". The frame of reference is all information contributing to the perception of the figure. Whereas the figure/ground constellation belongs to the phenomenal field, we are not necessarily aware of all the information in the frame of reference. Let me summarize:

1. The sensory field is the input to the sensory organs.
2. The perceptual field is made up by perceptual processes on any level of consciousness.
3. The phenomenal field is the perception of figure/ground.
4. The frame of reference consists of any information in the perceptual field and the mental field of relevance to the figure.
5. The figure/frame of reference constellation is the perspective.

If we could not change perspectives, we would be mentally locked in action. We would not be able to listen to the radio, while we are cleaning the house or follow a conversation, while we are driving. But these examples also show something else. The very fact that we can drive or clean the house while we focus on something else implies that although we do not attend to the ground, we can still perform in it accurately. This means that the background is preattentively perceived and acted on. For instance we are able to walk and take part in a conversation at the same time. Even if the terrain is difficult, we do not have to think about this. We do not take conscious decisions about how to compensate for up-hill/downhill variations, about stepping over a stone, about avoiding colliding with meeting people, or about following the turns of the path. If we are totally preoccupied with our thoughts, we make slight mistakes because we act totally absentminded (Where are my glasses?). But we can implicitly feel hungry, make a meal and eat it while talking about perception-action theory.

Applied to the fingerprint allegory above, this means that the police, according to the situation, choose a program (choose a perspective) which selects relevant information (frame of reference) but at the same time there may be other programs running in the computer such as downloading new info from the internet, scanning for virus, sending and receiving emails etc. These background activities have nothing to do with the search for the fingerprint match.

As long as these background processes proceed as expected, we do not have to attend to them. It is the mismatch that attracts attention. We are alerted (a pop up window on the screen of consciousness). This way our attention is directed to unexpected events. We are mentally where the testing takes place. This is the vital place for awareness to be.

Perception is a simultaneous capacity. As we perceive something consciously, we are able to perform various tasks – walk, wash up the dishes, windsurf, ski, use the right muscles for talking, and even play music. As can be seen from these examples, we have a capacity to perform procedurally without paying attention to it. Since we are able to act, we must have perceived environmental information to act on. These perceptions may not be the figure of the perception and it does not even have to be ground if it has no relevance to the figure. Still it is perceived; otherwise we would not be able to react to mismatches.

Now we have four levels of perception:

1. Figure (phenomenally attended object).
2. Ground (phenomenally conscious but unattended space in which the figure is situated).

3. Frame of reference (the ground + any information affecting the perception of the figure).

4. The perceptual field (the figure, the frame of reference + information of no relevance to the figure, which might affect the subject and guide procedural action).

This is a rough sketch, but it serves to help us in the analysis of music perception. As we shift perspective the perceived information takes new roles in this schema.

Perspective & time

Can time be perceived? Is not time a cultural invention, an abstraction from movement? The fact is that we do perceive duration. We react to deviations of expected durations (Mustovic *et al.* 2003)[100]. It is just the name for this that is cultural.

The perception of time is perspective dependent. The main distinction is between the subjective perception of the present moment and the cultural understanding of time.

There is an intentional aspect of perception, the prediction of the situation, which in turn entails direction – a sense of temporality. Edmund Husserl labelled this *the intentional arc* (Husserl 1969)[101]. The present moment is not just the edge between the past and the future. Husserl argued that the perceived now contains *retention* (the immediate past), *presentation* (the edge), and *protention* (the immediate expectation of the outcome of the situation). Daniel Stern (2004)[102] calls this *the present moment*. It is the time span framing perception. This is important because in perception-action theory, perception is treated in this sense – as the present moment experi-

[100] Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J.G., Hennig, J. & Seifritz, E. (2003) Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. *Neuroimage*. Vol. 20. No 2. 429–434

[101] Husserl, E. (1991) *On the Phenomenology of the Consciousness of Internal Time*. Kluwer, Dordrecht

[102] Stern, D. (2004) *The Present Moment*. W.W. Norton & Company, New York, London.

ence. This includes an experience of what we just lived and an anticipation and intention based on what just happened. Since the perceived moment has some expansion in time, we can perceive movements and actions – life as a process. The usual span of the present moment has been estimated to 2–4 seconds – the length of a breath and thus the length of a phrase. This time span is our window to life. It is the dividing line between experiencing the world directly ("live" and "on-line") and indirectly as the world symbolized, talked about or thought about in linguistic terms (the cultural perspective). There is a difference between the experience of living for five minutes as a moment-to-moment succession and the experience of thinking about the duration of five minutes. The first is phenomenal and the second is cognitive. These are two perspectives of time.

When challenged in physical activity that demands full concentration, such as playing tennis, we have to focus on the present moment. We merge with the activity. It is a here and now experience. This state has been labelled *flow* to denote a "state in which people are so involved that nothing else matters" (Csikszentmihaly 1990: 4)[103]. Examples of activities generating flow are play, art, ritual, pageantry, sports, music and making love. The self (the conscious thoughts of the self) is forlorn and the subject merges with the action-perception (the music, the mountain climbing, the computer game etc.). With the reservation for the highly intellectual activity of chess playing, mentioned as an example by Csikszentmihaly, I would say that most of his descriptions fit the descriptions of procedural perception-action activities. We seek these activities because of the pleasure to get totally absorbed which comes from focusing on the challenge at hand. We lose ourselves in a task that challenges our skills. We merge with the perception and the flow of time. It is the pleasure of the football player to juggle the ball, of a long distance skier to consume the landscape in five giant steps, of the windsurfer to plane, or the flamenco guitarist to execute an explosive *farruca.*

---

[103] Csikszentmihályi, M. (1990) *Flow – the psychology of optimal experience.* Harper & Row Publishers, New York

Perspective & movement

If you sit on a train at a railway station and another train passes the window, you may feel that your train is moving. However, when the train has passed the window, you realize that the station is still there. Now you change the frame of reference. This decides who is moving. You are tossed from a speed of 60 km an hour, say, to zero in a wink. This is a concrete example of perspective dependent perception of movement. How do our five perspectives affect the perception of movement?

1. Egocentric perspective: If I move I will bring the frame of reference with me. This will make me invariant and the surrounding variant. This is less obvious if I walk, and more obvious if I go by car. This is the perspective James Gibson is investigating in *The Ecological Approach to Visual Perception* (Gibson 1986).

2. In the allocentric perspective the landmark is allocated to the church, say. This means that I perceive myself as variant (moving) in an *invariant* (static) structure. The church has the same size wherever I go.

3. In the dyadic perspective I have a tendency to imitate – to synchronize my own movement with those of the other. As we walk side by side, I tend to adjust to your walking rhythm automatically.

4. In the tribal perspective I perceive myself as one of the members of a group. My movements and my intentions and my actions are integrated (conjoined) in this group activity. I tend to contribute to the achievement of common goals.

5. In the cultural perspective movement is something thought or talked about.

The fact that movement is perceived differently from perspective to perspective in combination with the depiction of music as a sensory-motor schema supports the idea of a perspective theory for music perception.

Perspective & music

Since the perspective is critical to the perception of movement and time, it is critical to the perception of music. If we did not take perspectives in music, we would be swamped by the flow of information. Janata & Grafton (2003)[104] have estimated that a music passage may consist of hundreds of thousands of elements that can be readily memorized and remembered. Since we cannot attend to more than a tiny fraction of this information at the same time, the brain has to choose a perspective.

The perception of music usually refers to the case when music is in the focus of attention (figure), but I think that the other levels of perception – music as ground, music in the frame of reference and music in the perceptual field – are worth consideration:

1. Music as figure of perception. This is the fraction of music that we are capable of attending to (usually the melody). In this case we can apply the five perspectives listed above.

2. Music as ground. This is everything in the sounding music that we do not attend to, but are phenomenally aware of. We have two cases: A) The object attended to is musical, i.e. the harmony and the rhythm as ground to the melody. B) The object attended to is not musical. This is in a sense to perceive the world as if it was music. Examples: Silence in a musical context is a pause. Music could animate the landscape: "The Hills are Alive with the Sound of Music". Since our concern in this investigation is music perception and not about the perception of non-musical objects, this option (2B) will not be further discussed.

---

[104] Janata, P. & Grafton, S.T. (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*. Vol. 6. 682–687

3. Music in the mental field (internal models, musical memories, familiarity with the style). This creates expectations in the musical development.

4. Music in the perceptual field. This is when the music does not have relevance to the attended task at hand i.e. reading a book with background music. But we are still inclined to interact with it. ERP experiments (c.f. p. 38) show that music can be pre-attentively perceived (as when we read a book and wiggle the foot rhythmically to music).

The four options in the model can be reduced to three. Since level 2 and 3 are both frames of reference, we could state that music can be figure, frame of reference, or irrelevant to the perception of the figure. The last case demands that the figure is not musical.

Do we really interact with unattended music? Yes, it is even possible to sing and play and think about something else. I know this because I have experienced that I can be distracted by the audience without getting lost in the performance. The knowledge how to play is, once it is learned, procedural. It is the knowledge of the body. Therefore it is not at odds with thinking. I have seen musicians who go on playing while they answer complex questions. The performance might even in some respects improve when the music is out of focus:

> In the conscious movement, the velocity is slower and accuracy is poorer. This means that you have lost the on-line control of the movement, which works automatically: instead, you have used a system that is slower and is not adapted to doing automatic movement. (Gallagher & Jeannerod 2002)[105]

We do not necessarily bike or swim better if we concentrate on how to do it. The same can be observed in music, the autopilot takes you home. It is well known among musicians that too much thinking about what you are doing makes you lose timing.

---

[105] Gallagher, S. & Jeannerod, M. (2002) From action to interaction. *Journal of Consciousness Studies*. Vol. 9. No 1. 3–26

How should we describe music in the perceptual field? Imagine that you are fully concentrated on a speech you are going to present, while there is music in the background. The music has no relevance to the speech. Even in this case it can be observed that you move to the music. You perceive it but you are not aware of it. However if something unexpected occurred in the music such as a false tone or an unexpected chord, this could draw your attention. The fact that such mismatches can make you attend to the music proves that you did perceive it on a pre-attentive level.

There are several possible ways of hearing movement in music. You can hear the sound source moving, as when the ice cream van passes you playing its tune. The same thing happens if the sound is moving from loudspeaker to loudspeaker. This is movement relative to the own body. It is an egocentric experience.

But, since music changes over time, there is an experience of movement even if the source is stationary. There are two options: Either you perceive yourself as stationary and the music as moving, or you perceive yourself as moving and the music as stationary.

The perspective theory of music perception

I have tried to explain the difference between the sensory input and the perception of music. The former is not consciously perceived. Perception is dependent on the meaning provided by the perspective. Through genetic factors and learning we tend to perceive holistically rather than particularly. If we have the information needed to perceive holistically, we will do so and the detail will be perceived in that context. In music we perceive relations, Gestalts and movement. Procedural learning and perceptual learning is important in this process.

What we consciously perceive is not the sensory input but what we make of it. Imagery is an important agent in this transformation. I am using a perception-action understanding of imagery, where imagery is a continuum from perception, via complementary imagery to clear-cut imagery. The perception process implies the following stages:

1. The *situation* (the sensory field) brings on a *perspective* [implicitly selecting task relevant information (*the frame of reference*)].
2. *Imagination* is to integrate the perspective selected information (the content) with an egocentric frame of reference. It is a "suggestion".
3. The suggestion presents itself as a form or a figure (a perception).
4. If I act on the perception the situation will change and the process becomes cyclical.

As can be seen from this, the theory states that perspective taking is pre-attentive, prior to imagination and prior to phenomenal perception.

To sum up: Music as figure of perception is charged with meaning. The meaning is dependent on the perspective. This leads to *a perspective theory of music perception:*

1. The egocentric perspective: the frame of reference is the listener's body. In this perspective musical movements are perceived relative to the body: A passing ice cream van, a passing military parade, music moving from one loud speaker to the other etc.
2. The allocentric perspective: the landmark is a point in the terrain. Does it have to be the terrain? Let us replace the world terrain with the more general word "structure" and ask: Can this perspective be applied on music? Let me exemplify this with tonal music. I suggest that the allocentric landmark is the tonic and that the sense of tonality serves as a frame of reference. Music is perceived as hierarchically organized around a reference tone (the tonic). All the other tones are related to this reference point (Krumhansl 1990)[106]. The tonic and the hierarchical surrounding space

[106] Krumhansl, C.L. (1990) *Cognitive Foundation of Musical Pitch.* Oxford University Press, New York

are implicitly perceived (Granot & Donchin 2002)[107]. The sense of tonality does not require musical training. From this landmark we can specify perceived locations in the music. Tonality allows chord progressions and melodies to lead to these allocentric landmarks which are felt as resting points in the music. In the allocentric perspective I perceive my own movement in a fixed music structure as if moving in a landscape. This is however not dependent on tonality. As soon as the music is organized along some principle that can be perceived the possibility of navigation in the music arises. This organizing principle is the frame of reference (the internal model). It is a set of musical rules. In the allocentric perspective the structure is static. But, when the allocentric perspective is integrated with the egocentric perspective I will perceive myself as moving in the music structure.

3. The dyadic perspective: The landmark is the body of the other and the frame of reference is information space shared by two people. When I integrate this perspective with my own I perceive the situation from the other's body. And feel the other's movement in my own body. This is for example to perceive music as the movement of the musician as if it was me moving.

4. The tribal perspective: the frame of reference is the information space shared by a group. Joint action experience. Playing in groups. In the audience, I will interact with the performing group. So will the other individuals in the audience. Thus the band will create an interacting crowd. We take part of a ritual acting out movements together. When I integrate this perspective with my own perspective, I perceive myself and the world through the eyes of the group.

---

[107] Granot, R. & Donchin, E. (2002) Do Re Mi Fa So La Ti – constraints, congruity, and musical training: an event related brain potentials study of musical expectancies. *Music Perception*. Summer. Vol.19. No 4. 487–528

5. The cultural perspective: the frame of reference is declarative (coded) knowledge available in the culture (context). The music is a sign of cultural knowledge. Such contextual knowledge about the composer, say, can be integrated in my perception of the music. For example, the declarative knowledge that Eva Cassidy died from cancer, just after having recorded *Over the Rainbow*[108] affects my perception of the song. It charges the words with additional meaning and this affects my perception of Eva's voice and phrasing.

This is a schematic presentation. It is just to give the reader a first idea. The depiction of the perspectives, the consequences of the perspectives for perception and the integration of the perspectives will be elaborated in the following chapters. The listener alternates between these perspectives.

The principal motivation for this theory is thus that perspective shapes content to form. There is no representation in the four first perspectives. The form does not represent the content; it is the content. But, if the form is taken to mean something for the beholder, we have a representation – a sign. I will argue that only the cultural perspective permits semiotic analysis and that we have four additional perspectives where musical meaning is perceived directly.

[108] Eva Cassidy, CD: Over the Rainbow. *Songbird.* Blixtnet

2

Sound

The Egocentric Perspective

Depiction of perspective

The landmark in the egocentric perspective is the own body. The frame of reference for directions is the body schema, which involves the vertical and the horizontal through the body. From this landmark we perceive the direction and the distance to the sound source. It is the perception of the *auditory scene* (Mather 2006[109]; Bregman 1990).

We detect horizontal directions by interaural time difference. Depth is detected by a decrease in perceived volume. But the perception of depth also has to do with the fact that the environmental surfaces reflect sounds more if the source is distant. Thus close sound sources are perceived as clear and dry (no reverb) and vice versa.

The time frame is the present moment consisting of the sensory buffer of the immediate past and the anticipation of what will occur next. This allows for short sequences to be heard.

There is also a frame of reference for sounds consisting of internal models of sounds made up by memories of sounds and Gestalts of sounds.

The egocentric perspective is about the identification of the source, the direction to the source and the movement of the source.

---

[109] Mather, G. (2006) *Foundations of Perception*. Psychology Press, Hove

Thus this comes close to the realm of acoustics and particularly to the branch within acoustics specifying human responses to sounds.

The perception of sound has implicit meaning and sounds can affect us emotionally. Sound, in the sense of timbre is sadly neglected in musicology, considering its importance in all music (Wicke 2003)[110]. Maybe the reason is that music does not begin until the sounds are connected to structures. However, sound involves the production and tuning of instruments, the player's technique, orchestration and arranging, the sound engineering technique – in other words, the whole process of music production. It has become increasingly important for the identity of styles and groups. The sound colours the whole experience of music.

The perception of sound

Auditory discrimination is improved by perceptual learning (Wright & Fitzgerald 2003)[111]. Perceptual learning has two important aspects. The first is that perceptual learning is to learn directly and automatically from the perception. The second aspect is that the knowledge learnt is the knowledge to perceive better.

Perceptual learning is a pre-attentive adaptation to the environment (Gibson, E. 1953). Eleanor Gibson was opposed to the dominant nativism of Gestalt theory. I will however use perceptual learning and Gestalt complementarily. To some extent our tendency to perceive certain Gestalts must have been learnt through perceptual learning; to some extent it is innate.

One category of perceptual learning, discussed by Eleanor Gibson is *differentiation* – an ability improved by exposure to stimuli. This is not difficult to spot in music. The professional sound technician is the ultimate example of excellent music discrimination. To see a sound engineer at work is magical. Everyone who has been in-

[110] Wicke, P. (2003) Popmusic in der Analyse. *Acta Musicologica*. Vol. 75. No 1. 107–126

[111] Wright, B.A. & Fitzgerald, M. B. (2003) Sound – discrimination learning and auditory displays**.** *Proceedings of the 2003 International Conference on Auditory Display,* Boston, MA, 6–9 July. 228–232

volved in a recording and has experienced a sound engineer at work knows what a difference she can make. Everything she hears is connected to the controls on the sound board. With quick moves she is able to change the recording from a complete disappointment to the musicians to something they can be proud of. There are innumerable parameters to juggle: every instrument is picked up by one or several microphones, each leading to a channel, each channel is treated separately on volume, overtone register, reverb, sound effects, filters cutting frequencies and so on. A virtual acoustic scene is created through panning, creation of depth, and the separation of the soloist from the background. Now you can hear where the players are situated in your living room as you listen to the music. The first ambition is to get a clear picture that sounds nice. But in addition to this comes the sound ideal of the particular style and the particular group. An important part of this is the sampling technique – the recycling of loops from old recordings.

To do all this, the sound engineer must have learned to differentiate features in music unperceivable to others. This ability comes from the daily work, or "exposure to stimuli", as Gibson said.

A related category of perceptual learning is *stimulus imprinting*. Imprinting is exposure to a stimulus resulting in representations (neuron connection potentials). Through imprinting, neural structures are developed to specialize for a certain stimulus (Goldstone 1998)[112]. The effect of this is recognition and familiarity with that stimulus. This is how we learn to recognize sounds. *Feature imprinting* is the ability to select a crucial feature important to identification. As can easily be understood this mechanism has evolutional importance. It is vital to children to be able to identify parents by their voices. We have an enormous capacity to recognize voices.

If a singer has a characteristically deviant style of singing, she will stand out and her songs will stick in our memory. Louis Armstrong's singing is from the aspect of sound more deviant than his trumpet playing and thus, although the trumpet playing was excellent, and although he was originally an instrumentalist, the voice is more remembered. If a singer has a unique profile of voice or phrasing

---

[112] Goldstone, R. (1998) Perceptual learning. *Annual Review of Psychology*. February, Vol. 49. 585–612

(Stevie Wonder, M. A. Numinen, Bob Dylan, Björk, Billy Holliday, Mrs Miller) he/she has an advantage in his ability to reach out to a mass audience.

Pop music makes good use of stimulus imprinting of sound. The subtlest reason that pop music is so flavourful to our brains is that it relies strongly on timbre (Levitin 2007)[113]. Sound is an identification factor for styles, groups, and even particular songs. Popular performers or groups, Levitin argues, are pleasing not because of any particular virtuosity, but because they create an overall timbre that remains consistent from song to song. Groups have identifiable sounds. We know from a fraction of listening what it is. The sound works as a signal or a trademark: Here I am again, the music says. Just think of guitar riff in the intro of *Satisfaction* by the Rolling Stones. You can identify the tune from the first chord.

Visual parts that co-occur in a set of patterns will tend to be bound together. This is when one part makes us perceive the whole pattern. Through unitization (or pattern completion) (see Chapter 1), we tend to see wholes instead of parts, situations instead of details, affordances instead of objects etc. The implicit meaning is induced. The brain, as we have seen from the neurological evidence, automatically combines relevant information to create a situation understanding. In music we perceive sounds as their sources (the instruments, or the singers). The sound of an acoustic guitar may act as a cue to perceptive memories in all modalities of a guitar such as shape and colour, and even the smell of spruce wood, as well as the feeling of holding it and playing it, not to mention all the memories from former listening. This unified perception, multi-modal in character, creates invariance. The sound of a Les Paul guitar holds a universe of implicit understanding for the jazz guitarist.

When famous artists sing in medleys such as *We are the world*, each artist is imposing his or her character by means of voice to the extent that you feel: "Oh it is a Stevie Wonder song, ah no it is a Bruce Springfield song, ah no Bob Dylan must have written it" and so on. Thus the perception of the composition changes with each artist just because of this one feature. Styles can be detected from the mere sound of a typical instrument such as the wa-wa-guitar

[113] Levitin, D.J. (2007) http://www.iht.com/articles/2007/01/03/features/music.php?

pointing to funk music, Les Paul guitars for jazz, Spanish guitars for flamenco etc. And, if sounds can be associated with music styles, they can also be associated with music tribes (the bebop tribe, the metallica tribe, the indie tribe etc.)

These unifying effects should be distinguished from our ability to perceive Gestalts. Gestalts are not in the same sense resulting from pattern completion. Gestalt mechanisms are certainly unifying too. We perceive timbres rather than the separate frequencies in the overtone spectra, chords rather than the separate tones which make up the chords, and orchestral sounds rather than the separate instruments producing this sound, and melodies rather than isolated tones.

Egocentric emotions

In this context "egocentric emotions" does not mean: I am the most important person in the world, but merely emotions emanating from egocentric perspective perception. Behaviourist theory of emotion is useful in the egocentric perspective context. This is a basic understanding of emotion. It does not demand higher order perception. Edmund T. Rolls states:

> Emotions can be defined as states elicited by rewards and punishers. (Rolls 2005: 11)[114]

A reward is anything for which an animal will work and a punisher is anything it will avoid. Rewards and punishers are called reinforcers because they reinforce animal behaviour. Rolls distinguishes primary reinforcers from secondary reinforcers. Primary reinforcers, such as food and sex and snakes, are gene specified reinforcers. The genes specify the goals for the action. This means that there is a fixed response to stimuli. Secondary reinforcers are objects that we have implicitly learned to associate with primary reinforcers. The reward/punishment is an operant in this learning. This is for example action/outcome learning.

---

114 Rolls, E.T. (2006) Brain mechanisms of emotion and decision taking. *International Congress Series.* Vol. 1291. June. 3–13

How can we think of music as a primary reinforcer? Any sound loud enough is frightening. The amygdala (usually and primarily associated with fear reactions) seems to react to sound intensity (volume and pitch) more than to identity (LeDoux 2003). Thus loud sounds are, as we all know, frightening irrespective of source.

In addition to that it seems that the quality of the sound has emotional valence. One of the very first brain scanning tests for emotional responses to music (Blood *et al.* 1999)[115] examined brain activation for consonant/dissonant chords. The evidence that orbitofrontal cortex and the limbic system were activated indicates that dissonances and consonances are emotionally felt. In other words, this confirms what we already know – there are pleasant sounds and there are unpleasant sounds. The study also showed that this was not the result of any conscious understanding of the perception.

This can be compared with the perception of surfaces. Some things are nice to touch because they are soft and some are unpleasant because they are uneven. Sounds are thus in themselves rewarding or punishing. They are primary reinforcers. This means that we do not explicitly (consciously) judge the sound to be pleasant or unpleasant; we experience this directly.

In a study of emotion processing in passive listening to major, minor and dissonant chords (Pallesen *et al.* 2005)[116] found that the amygdala, the retrosplenial cortex (BA26), the brainstem and the cerebellum were critical areas. The authors conclude:

1. The areas were activated even by single chords.
2. Emotional processing is enhanced even in the absence of cognitive requirements.
3. Musicians and non musicians do not differ.

---

115 Blood, A., 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.* Vol. 2. No 4. 382–387

116 Pallesen, K.J., Brattico, E., Bailey, C., Korvenoja, A., Koivisto, J., Gjedde, A. & Carlson, S. (2005) Emotion processing of major, minor and dissonant chords: an fMRI study. *The Neurosciences of Music. Annals of the New York Academy of Science.* Vol. 1060. 450–453

An example of a primary reinforcer is the music in the murder scene in Hitchcock's Psycho. The sound is produced by sharp violin attacks in the top register. This has an alarming effect on the listener. It creates fear, panic even, which is bodily and acute. Many alarms sound somewhat like this and they cannot be misinterpreted. The sound is dissonant, high in pitch, loud and repeated.

There is an interesting relation between harmony in music and environmental sounds. Any sound consists of a combination of overtone frequencies (a spectrum). When a string is hit, not only the full length of the string will vibrate, but also several subdivisions of the string length. These subdivisions produce subdivisions of sound waves. The length of the wave is related to the frequency. Thus when a string is hit, we do not just hear the frequency of the full length string but all the frequencies emanating from ratios of the string. These sound waves are the overtones. Instruments are designed to pick up different overtone spectra of frequencies (resonance). The resonance is heard as the timbre of the instrument. Timbre is thus a spectrum of overtones.

If a vibrating string is divided, this results in the double frequency – the octave. The simple ratios were discovered by Pythagoras. They are 1/1 string length (prime), 1/2 (octave), 1/3 perfect fifth, 1/4 (octave no 2), 1/5 (major third). A triad consisting of the prime, the major (or minor) third, and the perfect fifth is considered consonant. These tones thus happen to be the first (and strongest) tones in the overtone series. Higher up in the overtone series, and much weaker than the octave, the fifth and the major third, we find the smaller fractions. These intervals are used in chords as colourings and as lead tones. They make the chord sound dissonant. Consonant chords sound pure, smooth and clear. It is a pleasant sound. Dissonant chords are perceived in the opposite way (unpleasant, unclean, rough etc.)

As any musician knows, we tend to get used to dissonant chords. A major 7^{th} chord is not dissonant to a practicing musician today. If a chord is perceived as dissonant or not consonant, depends on the listener, the era, the style, and the harmonic context. A chord might sound dissonant in one context but consonant in another. This has to do with the function of the chord and how the voices in the preceding chords prepare for the chord. If the "dissonance" is caused by rule-governed voicing it will not be unpleasant and thus not so dissonant. The same chord in another context might sound terrible. There is thus a top-down processing of harmony.

But in this perspective we are not making these considerations. We are only interested in the isolated sound impact. The point here is that the tones in consonant chords are the strongest frequencies in most natural sounds and voices. There is thus an expectancy factor in consonance, whereas dissonant sounds are unexpected and thus more alarming. In a dissonant chord we hear the distant overtones (the seconds) much louder than in normal environmental sounds.

"Dissonance" can be defined by its tone content. However it is also a depiction of phenomenal experience. This pleasantness/un–pleasantness is how the amateur would perceive consonance/dissonance. Thus we do not have to be aware what tones the chord consists of to tell if it is consonant or dissonant. Blood's study shows that we feel the difference between consonance and dissonance on a direct emotional level (Blood *et al.* 1999). We detect this as directly as we detect the difference between a smooth and a rough surface.

Sounds can also be secondary reinforcers. Since the unitization effect makes us perceive sounds, not just as sounds but as holistic multi modal situation understandings, there is a meaning right in the perception that can be emotionally experienced. The mere sound of the guitar creates emotions for guitarists and afinados because this sound is linked (pattern completion) to pleasant experiences. Music is, as are tastes and smells, a trigger of *episodic memories*. A well known episodic memory from literature is the taste of a piece of Madeleine biscuit soaked in lime blossom tea in Marcel Proust's *Remembrance of Things Past*. This taste was the cue that threw the author back to childhood. The whole situation including other senses, intentions and emotions is actualized – as if we were there. A short sequence of entire situations including the emotional aspect is replayed. For music this means that a tune could trigger an imagination of a place where it was first heard. We live the event anew. The phenomenon has been labelled the "Darling they are playing our tune – dimension" (Davies 1978)[117].

[117] Davies, J.B. (1978) *The Psychology of Music*. Huthinson, London

Karin Johannisson gives a range of examples of how strong such musically evoked nostalgic reminiscences can be. Since music is not verbal it can go beyond cognitive memories and connect the listener to her pre-linguistic world as a child – to “a space of innocence” – (Johannisson 2001)[118]. This space is an undivided whole, a world in which the child does not ascribe emotions to itself but perceives the world as emotional and alive. It is a magical world. It is in a Freudian sense The Mother.

I will leave this theme now since it is speculative, but my personal opinion is that music really can evoke strong oceanic longing to a paradise lost. We will come back to this in the final discussion.

Major/minor

Although many authors have written about the perception of major/minor chords and scales (Hindemith 1941[119]; Deutsch 1999[120]), it still remains a riddle why major chords are felt has happy and minor chords are felt as sad. The fact that the same brain areas (the amygdala, the retrosplenial cortex, the brainstem and the cerebellum) are critical to major/minor as well as to dissonance (Pallesen 2005), suggests that minor chords are perceived as slightly dissonant.

The minor chord, named so because the major third is replaced by a minor third, could be said to be somewhat more dissonant just because the minor third appears later in the overtone series than the major third. This alone makes the minor chord slightly more dissonant than the major chord, according to the principle suggested above that late tones in the overtone series appear as more odd. However the minor triad is not considered dissonant among musicians. I think there is an additional factor that might be considered. If you play an A minor triad – A, C, E – an additional C# will arise because C# is an early overtone to the fundamental A. There are chords where both the major and the minor thirds are played. In

118 Johannisson, K. (2001) *Nostalgia – en känslas historia.* Albert Bonnier, Stockholm

119 Hindemith, P. (1941) *The Craft of Musical Composition.* Associated Music Publishers, New York

120 Deutsch, D. (1999) *The psychology of Music.* Academic Press, San Diego

these cases the minor third is on top (higher in pitch) as a + 9 colouring of a dominant chord (e.g. E7+9 could be E, G#, B, D, G). These are clearly dissonant chords. But the case we are discussing is a chord where the major third would be on top of the minor third. This is too horrible to exist in common harmony. This dissonance can hardly consciously be heard, because the major third resonance is too weak for identification, but it is there in all minor chords and it cannot be ruled out that we react to this on a pre-attentive level. After all, the processing of consonance/dissonance according to Blood is pre-attentive. There is thus a possibility that although we cannot hear dissonance in the minor chord, we might react to an extreme dissonance strong enough to be perceived but to weak to be consciously identified.

Even if we for one reason or other react directly to minor and major, it is not as simple as that. Minor chords are not always sad. Dm chords and melodies in a C major key are perceived as using the C major scale and are thus not as sad as Dm chords in a Dm key which are perceived as using the Dm scale. Although the chords are identical, the tonal reference differs. It all depends on where the tonic is perceived to be. There are numerous variants of this but the simple rule is that the musical context affects the direct emotional perception of the chord.

Conclusion

The mere musical sound triggers emotions in the listener. It does so as a primary reinforcer (loud sounds, consonant or dissonant sounds), and as secondary reinforcers depending on the unitization effect in perceptual learning which connects the sound to implicit associations or explicit associations such as episodic memories. Most emotional reactions to sound are direct and pre-attentive. Such emotional reactions on the sound cannot be explained by the appraisal theory of emotion which suggests that we judge the perception and then react to it, unless we allow judgement to mean implicit judgement (see Chapter 1). Rather it is the other way around, we react directly emotionally to the sensory input and this emotion colours the perception of the sound as smooth, uneven, pleasant, unpleasant, calming, alarming etc.

This means that the figure in this perspective is not relevant. We react emotionally to sounds even if the sound is not the focus of our attention. It is the pre-attentive processing, the process that might or might not lead to attended perception of sound that produces the emotion.

3

Listening to the Other

The dyadic perspective

Mirror neurons make us copy the movement of the other "on-line". They are supposedly involved in the understanding of the other's intentions and emotions. In 2004 Helge Malmgren and I proposed that mirror neuron theory explains why music moves us physically and emotionally (Vickhoff & Malmgren 2005).

Recently there has been an accelerated activity in neuro-imaging the brain to study the connection between music and mirror neuron activity (cf. Molnar-Szakacs & Overy 2006)[121]. In this chapter I will present new evidence and develop the theory further.

Depiction of perspective

To copy the movements of another human being, we need a common neural format (Gallese 2003)[122]. This format is *embodied simulation*. Such simulation is possible through mirror neuron reactions to the movements of the other activating a pre-existing body representation in the brain – the body schema. According to mirror neuron

[121] Molnar-Szakacs, I & Overy, K. (2006) Music and mirror neurons: From motion to e-motion. *Social Cognitive and Affective Neuroscience*. Vol. 1. No 3. 235–241

[122] Gallese, V. (2003b) The manifold nature of interpersonal relations: the quest for a common mechanism. Published on line by the *Philosophical Transactions Royal Society*

theory, we automatically copy the movements of the other. We do this on-line with the other – in synchronization with the other. This has philosophical implications outlined by Metzinger & Gallese (2003)[123]. The authors distinguish a phenomenal first person perspective, which allows for the representation of a phenomenal second person perspective. In other words the second person perspective is integrated in the first person perspective. This is *the dyadic perspective*.

The dyadic perspective is taken from the other and the frame of reference is the information space shared by two individuals.

In this chapter I will argue that we can perceive music from a dyadic perspective. This means that the music is perceived *as the other* – as the emotional expression of a musician, say, and that we take the perspective of this other being. How can we do that? I claim that we feel the movement in the music and that this movement affects us. A beautiful example of this is the fact that patients with Parkinson's disease not only walk better if someone walks along with them, but also when they walk to music, even if the music is just imagined (Sacks 1990: 282)[124].

The dyadic perspective entails that we synchronize our movements with the other's movement through immediate implicit imitation. However strange this may sound, it is easy to exemplify. Let us say that we are watching a high jumper on TV. We wish that she can make 2.04 m and win the Olympic gold medal. As she passes the bar, we find ourselves lifting a leg to make sure that *we* will go clear. The synchronization of movement shows that we in that moment are not in the sofa at home but actually up there with her. We share motivation, situation understanding and emotion with the jumper. We perceive the situation as if we were there. We integrate her perspective with our own. Thus we do not just mechanically imitate but the whole perception of the situation is shared. In one experiment a preverbal infant was watching an experimenter unsuccessfully trying to undo a dumbell-like object. When the child was given the opportunity, it immediately tried to and succeeded in un-

---

[123] Metzinger, T. & Gallese, V. (2003) The emergence of a shared action ontology: Building blocks for a theory. *Consciousness and Cognition*. Vol. 12. 549–571

[124] Sacks, O. (1990) *Awakenings*. Harper Perennial, New York

doing it and seemed pleased at the result (Stern 2004). This is obviously more than just imitation. The whole intentional complex of motivation and emotion is transferred. This is situation understanding. It relies on implicit procedural knowledge and takes place without conscious effort. We automatically perceive movements as intentions related to goals. This is done by co-producing these movements/actions in our own body, living the intention of the other.

Do we take the dyadic perspective listening to music? Music triggers (as we all know) movements (Fraisse 1982)[125]. There is obviously some kind of imitation going on. When we listen to music from a dyadic perspective, the other (e.g. the musician) will be in the frame of reference. This means that the music will be perceived as the other. This other (artist, musician) is what we implicitly imitate.

Music moves us physically, but does this imply that it moves us emotionally? The phenomenological psychologist Franz From concluded from various experiments that when we have to describe a behaviour sequence, we generally do so by indicating a perception of some psychological state in the behaving person (From 1971)[126]. We perceive, he stated, "the sense", a mental factor – an intention, a purpose, a meaning.

Our case will however be stronger if movement triggers emotion. This is, as we will see, the case.

To sum up, the conditions for dyadic emotions in music are:

1. We can imitate (explicitly or implicitly)
2. We can imitate immediately
3. We take dyadic perspectives listening to music
4. Movement triggers emotions

In the following I will discuss these conditions one by one.

---

[125] Fraisse, P. (1982) Rhythm and Tempo. *The Psychology of Music.* (Ed. D. Deutsch) Academic Press, New York. 149–180

[126] From, F. (1971) *Perception of other people.* Colombia University Press, New York

Implicit imitation

We can see how important imitation is in the animal world. The young imitate their parents and thus learn how to hunt, where to hide, where to find, what to fear and so on. Imitation can be voluntary, but in the case of animals we have to assume that there is no rational decision behind it. The fact that the offspring imitates must be ascribed to the ability to take the perspective of the parent. This is how the offspring learns how to cope with situations. The child maps the parent's body and body actions to reproduce them immediately or later in similar situations (Blandin *et al.* 1999)[127]. This is how procedural knowledge is acquired.

Babies imitate their caretakers. There is evidence of such imitation tendencies just minutes after birth. Neonates (30 minutes!) have been proven to imitate mouth movements (Meltzoff & Moore 1987)[128]. Needless to say, the infant cannot think: "My mother is smiling. She must be using her mouth muscles. I want to respond. Now, where are my mouth muscles?" It is thus reasonable to assume that we are all born with the capacity to feel the other's face in our own.

Imitation is the first sign of dialogue. The early pre-linguistic communication between child and caretaker is called *protoconversation*. This conversation is an exchange of facial expressions and, when vocal, it is the melody (prosody) of speech that is communicated. For this reason researchers on child development refer to this early exchange as a duet (Stern 1985[129]; Trevarthen 1992[130]). This duet develops emotional bonds between child and caretaker. The capacity of implicit imitation thus is vital in child development.

---

[127] Blandin, Y., Lhuisset, L. & Proteau, L. (1999) Cognitive processes underlying observational learning of motor skills. *Quarterly Journal of Experimental Psychology: Human experimental psychology,* 52(A). 957–979

[128] Meltzoff, A.N. & Moore, M.K. (1987) Imitation of facial and manual gestures by human neonates. *Science.* Vol. 198. 75–78

[129] Stern, D. (1985) *The Interpersonal World of the Infant.* Basic Books, New York,

[130] Trevarthen, C. (1992) An infant's motives for speaking and thinking in the culture. *The Dialogical Alternative.* (Ed. A.H. Wold.) Scandinavian University Press/Oxford University Press, Oslo and Oxford. 99–137

Musical elements participate in the process of communicative development very early. "They pave the way to linguistic capacities earlier than phonetic elements" (Papousek 1996)[131]. Trevarthen has found temporal characteristics communicated through body-imaging based on "intrinsic motive pulse" integrating attention, learning and self-regulating physiology. To a large extent the beginning of communication starts with a sharing of patterned time (Trevarthen 1999)[132]. Imitation is seen as a condition to start information exchange before any kind of language is developed. It is the basis of social interaction.

Studying patients with prefrontal lesions, Lhermitte found that these patients compulsively imitate gestures or even complex actions performed by an experimenter (Lhermitte *et al.* 1983)[133]. This behaviour has been explained as an impairment of the inhibitory control normally exerted by the prefrontal areas. This indicates that if we were not inhibited, we would imitate constantly. We always have an imitation tendency, but it is not always displayed (Baldissera *et al.* 2001)[134].

Other evidence for implicit imitation has been presented by Dimberg, Thunberg & Elmehed (2000)[135]. The authors measured the activity in face muscles for subjects watching pictures of smiling and angry faces. Facial mimicry reactions are measured by EMG (electromyographic activity). The pictures were presented for 30 milliseconds and were then replaced with neutral faces. The results show muscular activity after 50 ms for smile and anger expressions. Both the time for exposure of the target and the reaction time are too

---

131 Papousek, H. (1996) Musicality in infancy research: biological and cultural origins of early musicality. In *Musical Beginnings* (Eds. I. Delíege & M. Sloboda.) Oxford University Press, Oxford

132 Trevarthen, C. (1999) Musicality and intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae.* Special issue. 155–215

133 Lhermitte, F., Pillon, B. & Serdan, M. (1986) Human autonomy and the frontal lobes: Part 1. Imitation and utilization behaviour: A neurological study of 75 patients. *Annals of Neurology.* Vol. 19. 326–34

134 Baldissera, F., Cavallari, P., Craighero, L. & Fadiga, L. (2001). Modulation of spinal excitability during observation of hand actions in humans. *European Journal of Neuroscience.* Vol. 13. 190–194

135 Dimberg, U., Thunberg, M. & Elmehed, K. (2000) Unconscious facial reactions to emotional facial expressions. *American Psychological Society.* Vol.11. No 1. 86–89

short for conscious processing of the stimuli. The conclusion is that there is an immediate and pre-attentive muscle reaction imitating the faces on the target pictures. It has been shown that subjects rating high in empathy (the ability to have the emotion of the other) show stronger EMG reactions (Sonneby-Borgström 2002)[136].

Zatorre (2003)[137] has shown that music activates hand function representation for musicians. Convincing evidence of implicit imitation in music perception is provided by Lahav *et al.* (2005)[138]. Non-musicians were instructed to learn a simple piece. When presented with the piece again they showed involuntary finger movement tendencies (in a specially designed glove) for heard tones. One single learning session is enough to establish the sound-motor connection.

Many areas implicated in music performance and music perception overlap: Cerebellum (generally associated with motor commands) is activated by harmony, melody and rhythm (Zatorre 2003). Other examples are the temporal and frontoparietal regions (Koelsch *et al.* 2005[139]; Tillmann *et al.* 2003a[140]) including Broca's area (Lahav, Salzman & Schlaug 2007)[141].

The lateral pre-motor cortex is critical to perceptual sequencing (Schubotz *et al.* 2004)[142]. Even when musicians just imagine playing,

---

136 Sonneby-Borgström, M. (2002) Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology.* Vol. 43. No 5. 433–443

137 Zatorre, R.J. (2003) Neuronal specialization for tonal processing. In *The Cognitive Neuroscience of Music.* (Eds. I. Peretz & R.J. Zatorre) Oxford University Press, New York. 231–236

138 Lahav, A., Boulanger, A., Schlaug, G., & Saltzman, E. (2005) The power of listening: auditory-motor interactions in musical training. *Annals of the New York Academy of Sciences.* Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 189–194

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

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

141 Lahav, A., Saltzman, E. & Schlaug, G. (2007) Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience.* Vol. 27. 308–314

142 Schubotz, R.I., Sakreida, K., Tittgemeyer, M. & von Cramon, D.Y. (2004) Motor areas beyond motor performance: deficits in serial prediction following ventrolateral pre-motor lesions. *Neuropsycology.* Vol. 18. No 4. 638–645

the pre-motor cortex and the supplementary motor areas are activated (Meidell *et al.*)[143]; (Halpern 2001).

In singing there is evidence of interaction between a heard melody and the production of the same melody (Hickok *et al.* 2003)[144].

Mirror neurons

We can implicitly understand a situation and act accordingly. Neurological evidence for this is piling up. An important step was the discovery of so called *mirror neurons* in 1992 (Di Pellegrino *et al.* 1992)[145].This finding provided a neurological substrate for implicit imitation. Studies on monkeys revealed mirror neurons in the frontal lobe, so named since the very same neurons fired when the monkey was watching movements related to hand and mouth as when it performed the same movements.

When the monkey observes hand movements that are part of its repertoire, the seen action is automatically retrieved (without having to be executed) in parts of the motor system. Mirror neurons prepare us for imitating motor sequences (Rizzolatti *et al.* 2001)[146]. Imitating by means of mirror neurons is referred to as resonance behaviour.

These findings have caused extensive research on mirror neurons and in the prolongation the development of mirror neuron theory. This theory in turn can be classified as a perception action theory, since it states that we perceive the other through action (imitation).

---

143 Meidell, K., Gaab, N., Halpern, A. & Schlaug, G. Neural correlates of performance in a motor imagery task: an fMRI study with professional pianists. http://208.164.121.55/hbm2003/abstract/abstract1086.htm

144 Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. (2003) Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience.* Vol. 15. No 5. 673–682

145 Di Pellegrino, G., Fadiga L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992) Understanding motor events: a neurophysiological study. *Experimental Brain Research.* Vol. 91. No. 1. 176–180

146 Rizzolatti, G., Fogassi, L. & Gallese, V. (2001) Perspectives – neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews/Neuroscience.* Vol. 2. Sept. 667–668

There is strong evidence that mirror neurons exist in humans (for a review, see Rizzolatti & Craighero 2004)[147]. When imitation and observation tasks were compared, the researchers found a complete spatial overlap of activation between imitation and observation. The only difference was the magnitude of the response. The location is the "frontoparietal architecture for sensory-motor integration" (Rizzolatti net)[148].

Are there mirror neurons processing sounds? Single unit responses for listening and oral repetition have been found in the Sylvian fissure, the cleft at the parietal-temporal boundary between Broca's area and Wernicke's area (Hickok *et al.* 2003). The authors also found frontal areas that combined listening and rehearsal, consistent with the claim that auditory processing is part of a larger sensorimotor integration circuit.

In October 2003 the discovery of audiovisual mirror neurons was reported. They represent actions independently of whether these actions are performed, heard or seen (Keysers *et al.* 2003)[149].

These findings propelled several neuro-imaging studies on music and mirror neurons. Musical experience or expertise can modulate the activity within the fronto-parietal mirror neuron system (Haslinger *et al.* 2005)[150]. Dancing experience has the same effect (Cross *et al.* 2006)[151].

In an fMRI study (Lahav, Salzman & Schlaug 2007) non-musicians learned to play a non-familiar piece by ear. Then they were played the music (not performing any movements). This activated the areas associated with mirror neurons: the frontoparietal motor related network, including Broca's area (known to be involved in sequence

---

[147] Rizzolatti, G. & Craighero, L. (2004) The mirror-neuron system. *Annual Review of Neuroscience.* Vol. 27. 169–192

[148] Rizzolatti, G. interviewed by Maddalena Fabri Destro on the net. http://scholarpedia.org/article/talk:mirror_neurons

[149] Keysers C., Kohler E., Umiltà, M.A., Nanetti, L., Fogassi L. & Gallese V. (2003) Audiovisual mirror neurons and action recognition. *Experimental Brain Research.* Vol. 153. 628–636

[150] Haslinger, B., Erhard, P., Altenmuller, E., Schroeder, U., Boecker, H. & Ceballos–Baumann, A.O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience.* Vol. 17. 282–293

[151] Cross, E.S., Hamilton, A.F.D.C., & Grafton, S.T. (2006). Building a motor simulation de novo: observation of dance by dancers, *NeuroImage.* Vol. 31. 1257–1267

processing in speech and music), and the pre-motor region. A presentation of the notes in a different order activated the same network but to a lesser degree. Listening to motorically unknown music did not activate this network. The authors conclude that the findings support the hypothesis of a "hearing-doing" system comparable to the seeing-doing previously observed. This means that as we listen to music, we latently perform the movements required to play it, provided we have played the music before. Such a system, the authors speculate, "may have been developed to protect the survival of all hearing organisms, allowing the understanding of actions even when they cannot be observed but can only be heard (for example, the sound of footsteps in the dark)". In this experiment the neurons carry out the same function and have the same location as the mirror neurons in the first experiments.

In order to entrain the movement of the other it takes more than imitation. We have to be able to predict the movements of the other. Umiltà (*et al.* 2001)[152] showed that mirror neurons react to predictions. Seeing someone with the intention to grasp an object releases a neurological grasping reaction in the observer, even if the actual grasping was never observed. This was confirmed by evidence indicating that viewers not only imitate but even show motor activation in advance of the observed action (Kilner *et al.* 2004)[153].

This leads to the question: How do we perceive the intention and emotion of others? Mirror neuron activation would not suffice to explain this. We have to look for a circuit which implies the limbic system. It has bee suggested that the anterior insula mediates between mirror neuron areas and the limbic system. The frontoparietal circuit which is reported to have mirror neurons is connected to the insula which has a

---

[152] Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*. Vol. 32. 91–101

[153] Kilner, J.M., Vargas, C., Duval, S., Blakemore, S.J. & Sirigu, A. (2004), Motor activation prior to observation of a predicted movement. *Nature Neuroscience*. Vol. 7. 1299–1301

> [k]ey role in affect generation and serving as a relay between frontal cortices and limbic structures, thus representing a possible pathway for empathic resonance. (Iacobini and Lenzi 2002)[154]

In a beautifully designed fMRI experiment Iacobini and colleagues showed that pre-motor mirror neurons are involved in understanding the intention of others. This finding made the authors conclude:

> To ascribe an intention is to infer a forthcoming goal, and this is an operation that the motor system does automatically. (Iacobini *et al.* 2005)[155]

In 2006 evidence was attained that the insula is the interface between mirror neurons and the limbic system (Dapretto *et al.* 2006)[156]. The involvement of the insula is interesting since this area has been reported to be involved in music perception and in the so called "song circuit". We will come back to this. Since the insula connects the mirror neuron area with the limbic system, "the meaning of the imitated or observed emotion is directly felt and understood" (Dapretto *et al.* 2006). Thus we can understand this as a *mirror neuron system.*

A prerequisite for prediction is that the movement is simple, stereotyped or familiar (Wilson & Knoblich 2005)[157]. These conditions are largely fulfilled in musical movements. In fact the authors use music as one of their examples.

But there is one more condition for prediction and this is that the observed individual has a motor program similar to the observer

---

[154] Iacobini, M. & Lenzi, L.L. (2002) Mirror neurons, the insula, and empathy. *Behavioural and Brain Sciences.* Vol. 25. 39–40

[155] Iacobini, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C. & Rizzolatti, G. (2005) Grasping the intentions of others with ones own mirror neuron system. *PlosBiology.* March. Vol 3. No 3. 529–535

[156] Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y. & Iacoboni, M. (2006) Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience.* Vol. 9. No 1. January. 28–30

[157] Wilson, M. and Knoblich, G. (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletine.* Vol. 131. No. 3. 460–473

(Knoblich *et al.* 2002)[158]. On a rough level this means that it is easier to predict the movements of a conspecific, but on a more fine-grained level it means that it is easier to predict the movements of some people than of others. This can be due to similar training. A bicyclist is better in predicting the movements of another bicyclist, and a pianist can predict the movements of another pianist.

In 2007 we were reached by the sensational finding that that mirror neurons are more engaged in *complementary action* (how to contribute in a situation) as compared to imitation (Newman-Norlund *et al.* 2007)[159]. Perhaps "mirror neurons" is not the most adequate name for these neurons? The finding could explain *complementary emotions* such as fear as a response to rage and tenderness as a response for sadness. In fact it becomes increasingly evident that mirror neurons are implicated in a range of behaviours. Rizzolatti does not talk about one mirror neuron system but of several mirror neurons systems.

To sum up, the mirror neuron systems are involved in the following phenomena:

1. Imitation
2. Implicit understanding of the other's intentions
3. Situation understanding including goals
4. Prediction of movement
5. Entrainment
6. Complementary action

---

[158] Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*. Vol. 55. No 3. 1027–1046

[159] Newman-Norlund, R.D., van Shie, H.T., van Zuijlen, A. & Bekkering, H. (2007) The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*. Vol. 10. 817–818

Do we take dyadic perspectives listening to music?

Arguments for the connection between music and sociability can be found in literature studying the interest in music of patients suffering from syndromes affecting their social ability. Patients suffering from William's Syndrome, a genetically caused disorder characterized by strong social interest and empathy, are generally interested in music, while patients with Asperger's Syndrome (low social interest) are less inclined to music activity.

Patients suffering from Williams Syndrome have been described as mentally asymmetric. The frontal lobe, the temporal lobe and cerebellum are enlarged in the few brains examined so far. This is at the cost of other areas. I.Q. rates around 60 (Levitin 2005)[160]. This group show a relatively spared ability in four cognitive domains: face processing, sociability and empathy, language, and music. They show "exaggerated prosody", they tend to talk about musical instruments and, surprisingly, even their motor abilities seem to be musically domain specific. Some individuals can play the clarinet or piano but cannot button their shirts or tie their shoes.

Inquiries prove that people diagnosed with Williams Syndrome show high sensitivity to sounds, show earlier musical interest, spend more time playing and listening to music, and are more emotionally affected by music than others (Levitin 2005). The musical capacity proves to be average. From this we can conclude that this groups stands out musically not because of capacity but because of interest. They live in a musical-emotional world. The brain scanning revealed higher amygdala and cerebellum activation by music for WS individuals than for controls (Levitin *et al.* 2003)[161].

---

[160] Levitin, D.J. (2005) Musical behavior in a neurogenetic developmental disorder – evidence from Williams Syndrome. *Annals of the New York Academy of Sciences* 1060. *The Neurosciences and Music II: From Perception to Performance*. (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 325–334

[161] Levitin, D.J., Menon, V., Schmitt, J.E., Eliez, S., White, C.D., Glover, G.H., Kadis, J., Korenberg, J.R., Bellugi, U. & Reiss A.L. (2003) Neural correlates of auditory perception in Williams Syndrome: An fMRI Study, *NeuroImage*. Vol. 18. No. 1. 74–82

Because of the evident connection between sociability and music interest in this group, which can be contrasted to the disinterest in music for autistic people, it has been suggested that musical and social interest are genetically interlinked (Huron 2001a)[162].

What do we imitate in music? Let us break down this question into two parts: Do we imitate performers? And: When no performer is present, do we imitate the music *as if* a performer was implicated?

The first question can be answered affirmatively just from the fact that we imitate each other generally. But this can be developed in the case of music. Anthropological studies confirm what child psychologists have found. Music creates social bonds. This starts in infancy, not only by means of nursery songs, but by all the musical features in mother-child communication (Dissanayake 2000a)[163]. Music is, as a feature of communication, lived from a dyadic perspective from the very start.

In a concert situation, we see the musicians in action. Musicians show their interpretation to the audience with their bodies.

> The sight of the gestures and movements of the various parts of the body producing the music is fundamentally necessary if it is to be grasped in all its fullness. (Igor Stravinsky 1936: 122)[164]

> A musician cannot move others unless he too is moved…for revealing of his humour will stimulate a like humour in the listener…Those who maintain that all of this can be accomplished without gesture will retract their words, when owing to their insensibility, they find themselves obliged to sit like a statue before their instrument…fitting expressions help the listener to understand the meaning. (C. P. E. Bach 1949: 152)[165]

---

[162] Huron, D. (2001a) Is music an evolutionary adaptation? *The Biological Foundations of Music*, Vol. 930. (Eds. I. Peretz & R.J. Zatorre) *Annals of the New York Academy of Sciences*. 43–61

[163] Dissanayake, E. (2000) Antecedents of the temporal arts in early mother infant interaction. *The Origins of Music* (Eds. N. Wallin, B. Merker & S. Brown). The MIT Press, Cambridge, Mass. 389–410

[164] Stravinky, I. (1936) *Cronicle of my Life*. Victor Gollancz Ltd., London

[165] Bach, C. P. E. (1949) *Essay on the True Art of Playing Keyboard Instruments*. W. W. Norton & Co., Inc., New York

This communication could be intentional, but in most cases it is spontaneous. I say this from my experience as a performing musician.

We can see body communication between the conductor and the orchestra. The pronounced purpose of conducting is to show the musicians when and how to play, but anyone who has seen a maestro in action (e.g. Sergiu Celibidache) knows that this is just a part of it. The bodily expression is enormous. The unpronounced effect of this is that it inspires the musicians and shows them what this music emotionally is about. The conductor acts out the musical emotion with his bodily and this emotional expression in turn is played by the musicians.

This is just as important to the audience. In symphonic performances the conductor takes on the soloist's role to impersonate the music. The movements of the conductor are connected to the music. The movements of the music and the emotion of the music are in this case not only heard but *seen* by the audience.

Dancers have made it their profession to express music with their bodies. Most styles of music have their own body expression. In many cases the music has evolved from body movement patterns just as much as the movements have evolved from the music. It is certainly true in the case of contemporary music that the music may be composed to the choreography. The Pilobulus Dance Theatre often works this way. But I think this is in some sense true for any dance music. Some music becomes popular because it is nice to move to and this popularity will make other composers try to make music like that. So there will be a selection in an almost Darwinian sense that will favour the survival of the fittest (the music that fits bodily expression).

Artists capture their audiences with their expression. In popular music, particularly, we have the phenomenon of idolization – a strong identification force. We can see how audiences, as they become absorbed by the performance, tend to forget about themselves and, seemingly unaware, move their lips in accordance with the song.

* * *

Now let us address the case when there is no performer present. What happens when we listen to the radio or the CD player at home?

We have a frame of reference where music is connected to body movement from the type of experiences exemplified above. We have seen musicians, artists, conductors and dancers move with the music. If you have seen Elvis Presley in the movie *Jailhouse Rock,* you will most likely have a sense of his dancing moves whenever you listen to the song.

Now consider this: Even if music is not movement in a spatial sense, *something* is moving, namely the musician. Music is produced by bodies in motion. The sound mirrors the musician's movements vis-à-vis the instrument. Musicians move their entire bodies to sing, or to blow wind or brass instruments. In fact, if we were to invent a device to register in sound how a person is moving her fingers over a table, we would end up with something like the piano. The piano registers the attack, the direction (left-right), the intervallic leaps, number of fingers used, the spread of the fingers and the speed as sounds. Every subtle nuance in tempo and touch is heard. Every heard change in pitch and dynamics corresponds to a movement by the player. There is a determinate relation between music and body movement. This goes for any instrument including the voice. Thus we can state that music is, in a concrete sense, heard movement. Movement is immanent in the heard information. We hear the performer in action. In this respect there is no difference from seeing the musician; it is just another modality. It is just routine-thinking to presuppose that movement must be seen. We can hear someone walk over a floor. The dyadic perspective implies that we are imitating the musician rather than the music. This has consequences for the analysis of perception. Let me point to three implications:

1. In perception-action theory perceiving is interacting with or even enacting the object. As we have seen from the neurological evidence, we interact not only when we play or dance, but even when we listen to music. How can, one might wonder, we interact with something as abstract as music? And how can abstract sounds evoke human emotions? But consider that we in the dyadic perspective interact with the

musician rather than with the music. This suddenly makes the interaction concrete.

2. As discussed in the theoretical chapter situation understanding (Koffka's demand, Merleau-Ponty's telos, and Gibson's affordance) is the implicit meaning of perception. What is situation understanding in the perception of music? This problem is solved if we, again, consider that it is the musician rather than the music that is perceived. The dyadic perspective undoubtedly opens up for the possibility to perceive intentions, emotions and goals. Even if there is a difficulty in stating that we can perceive intentions in music, there is certainly no problem with perceiving the intention in a person expressing herself through music.

3. As we perceive the musician through the music, we experience her grace, perfection and good technique. Technique is about achieving a goal with a minimum of effort. In music this can be heard as a relaxed precision in timing and pitch. It is a world of difference between listening to a singer who can hit the tones and one who has difficulties. As we implicitly imitate the other her lack of technique becomes our trouble. We suffer when the singer has difficulties in hitting the tones. The skilled and talented musician makes this without effort. To experience this could in some sense be compared to watching the Brazilian soccer star Ronaldinho with a ball. It is pure joy.

Professional instrumentalists can listen to a tune and then play it. They are duplicating the movement pattern of the first performer just from the sounding music. *Music is choreography.*

Do non-musicians imitate this movement? Yes, to a certain extent. An obvious feature in music is that it propels movement. We jog, dance, run, pump weights, play air guitar, and mime in skipping-ropes to music (some of us). Rap music is largely a way to move, even to the extent that there is a fashion of lose cloths to go with this lax rhythmic pattern. It is quite hard to sing a song by a famous artist (e.g. karaoke) without impersonating the voice as well as the moves.

In fact one of the basic functions of music is to guide movement. Let me remind of the working songs of seamen (shanties), railroad workers, and the slaves.

So music makes us move, but is this really to be compared to choreography? We could expect that a dancer who has once learned to dance to a piece of music, would, in accordance to the experiments on musicians, latently perform the dance movements when she listens to the music. But this movement pattern would be different from the conductor's movement and the conductor's movement would be different from the movements of the musicians. And one musician does not move as the other does since the movements are adjusted to the instruments. Thus we cannot claim that there is general relation between music and movement. Music is choreography, but it is not the same choreography for every category of listeners.

However there is a certain resemblance between these different patterns. Directions are often sensed as the melody "moves" from *low* to *high* registers and vice versa. Timing and the intensity plus an amount of structural features such as legato, leaps, and syncopations – are felt and imitated.

To sum up: We can take dyadic perspectives as we listen to music. According to the theory presented in Chapter 2 this means that we integrate the perspective of the other in our egocentric perspective. It is *as if* we were playing. If the musician is present this is no problem. If the musician is not physically present her movements are still heard and can to some extent be imitated.

The persona and the metaphor theories sketched in the introductory chapter imply that the perception of music is associated with human movement. But what I suggest here is that what we perceive in the dyadic perspective is not the music, but the musician. This has consequences for the analysis of interaction, situation understanding and the perception of technique. The music is the expression of the other. Seen this way, we are confronted with a situation – someone is expressing herself intentionally and emotionally. The answer to the question, how something as abstract as music can move us, is: It is not the music but the musician that moves us.

Immediate imitation

Imitation can be overt or latent, immediate or delayed. Latent imitation, which always is immediate, can be seen only in the neuronal pattern (Rizzolatti *et al.* 1999)[166]. There is proof of immediate imitation of music (Lahav *et al.* 2007). I suggest that immediate latent imitation paves the way for delayed imitation. In other words, this is how imitation is learned.

Immediate imitation demands action prediction. The superior temporal sulcus (STS) seems to be involved in action prediction, and in particular in the updating of predictions after a violation has occurred (Saxe *et al.* 2004[167]; Greeze *et al.* 2004[168]; Bosbach *et al.* 2005[169]).

Haueisen & Knoesche (2001) performed neurological studies on pianists listening to music and found statistically significant increase of activity in primary motor cortex and the supplementary motor area. The authors found that finger-related motor cortex activation in response to a particular note in a familiar sequence occurred 300 ms *before* the note could be heard. There is thus a continuous motor anticipation and activation as the music unfolds. If this anticipation did not exist, it would be impossible to play together with other musicians.

But not only fellow musicians share the time dimension, so does the audience.

This means that we must make a specification: *We do not, as we listen to music, imitate the movements of the other, but the anticipation of the movement of the other.* The word imitation becomes problematic as we become totally synchronized in sequences. Perhaps entrainment is a better word for this?

---

[166] Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (1999) Resonance behaviors and mirror neurons. *Archives of Italian Biology.* May. 137 (2–3). 85–100

[167] Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D.I. & Kanwisher, N. (2004) A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia.* Vol. 42. No 11. 1435–1446

[168] Grèzes, J., Frith, C.D. & Passingham, R.E. (2004) Inferring false beliefs from the actions of oneself and others: an fMRI study, *Neuroimage.* Vol. 21. No 2. 744–750

[169] Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Understanding another's expectation from action: The role of peripheral sensation. *Nature Neuroscience.* Vol. 8. No 10. 1295–1297

Entrainment theory is in fact not only about how we adapt our movement to each other, but about any movement adjustment. A person tends to move her limbs in ratios of 1/1, 1/2, 1/3 or 2/3 (Port *et al.* 1996)[170]. This is referred to as *self-entrainment.* It is a possible explanation for the ratios in musical rhythms and for the discrete time values. This is how we naturally move our limbs when we play and dance. Rhythmically complex music like flamenco reveals that the skilled musician is able to coordinate highly complex ratios.

Models for entrainment started to emerge in the mid 90s (Large & Kolen 1994)[171], (McAuley 1995)[172]. Entrainment was described as a form of prediction that may help establish *what* information in a pattern should be processed, as well as *when* that processing should occur. When sensory rhythmic input stops, sustained cortical activity continues to be observed at the conditioned frequency. Also, if pulses of the rhythmic stimulus are left out, the entrained neurons will fill in the missing beats. This connects to the complementary imagery theory (see Chapter 3). The mechanism is not thought to depend on any particular kind of stimulus. Rather it is a general capacity of the central nervous system to reflect temporal patterns in the environment.

J. Devin McAuley states in his thesis *Perception of time as phase: toward an adaptive model of rhythmic pattern processing* (1995) that time has three levels:

1. The ordinal scale (the order of events)
2. The interval scale (duration between events)
3. Tempo – a pattern of intervals constituting the beat

---

[170] Port, R., Tajima, K., & Cummins, F. (1996) Self-entrainment in animal behaviour and human speech. *Online Proceedings of the 1996 Midwest Artificial Intelligence and Cognitive Science Conference.* http://www.cs.indiana.edu/event/maics96/Proceedings/Port/port/.html

[171] Large, E.W. & Kolen, J.F. (1994) Resonance and the perception of musical meter. *Connection Science.* Vol. 2. No. 6. 177–208

[172] McAuley, J.D. (1995) *Perception of Time as Phase: Toward an adaptive-oscillator model of rhythmic pattern processing.* Technical Report No. 151 (Doctoral dissertation), Indiana University Cognitive Science Program

McAuley distinguishes two kinds of intervals: Filled intervals are the intervals filled with a sounding tone, whereas empty intervals are the intervals between onsets of tones. McAuley focuses on the latter type. Empty durations create the theoretical problem that there is nothing to be perceived, yet they are perceived. We can perceive time itself as durations. I believe that these intervals are empty only in the sense that there is an empty duration between events as focused figures. However there is a flow of external and internal background events constituting a frame of reference where the focussed events are allocated. If we for example take the own body as a frame of reference, the duration heard can be compared to own movement even if this is just eye movements or breathing.

In his thesis McAuley developed an *adaptive oscillator* model of entrainment. The concept adaptive oscillator emanates from the idea of the coupling of oscillators, which anecdotally is associated with the Dutch physicist Christiaan Huygens who 1665 observed from his bedside during an illness that two clocks hanging side by side were synchronized in the motion of their pendulums. He assumed that the clocks "communicated" through vibrations in the wall. And, as he moved one of the clocks to another wall the effect ceased to appear.

McAuley:

> The adaptive oscillator internalizes a beat, retaining a memory of that beat after the pattern stops. This creates expectancy for the occurrence of future inputs. (McAuley 1995: 12)

In the case of "missing" inputs, as well as when the rhythmic pattern stops, the adaptive oscillator continues to predict future inputs. In essence, the adaptive oscillator internalizes a beat, retaining a memory of that beat after the pattern stops.

Unfortunately this does not mean that McAuley or Large & Kolen have come up with the final scientific explanation of entrainment. The theory is mainly a mathematical model that can make computers "swing". The oscillator models do not show how this can be explained on a neuronal level.

The most salient feature in popular music is the beat. We know exactly when the next beat is going to happen – creating this "now-trance", which is such a typical musical experience. All the other events are ratios of durations between these beats. We feel the sixteenth notes clicking, even in the breaks. Not approximately but exactly. Entrainment can be pre-attentive, but if we focus on the music, the effect can be a desymbolization, where the past and future is lost in the intensification of the present. Paradoxically, the sense of time is lost. We are in the present moment – in the very beat. It creates authenticity: The experience of the unique moment lived here and now. Sublime experiences in music are often reported as a loss of the sense of time (Gabrielsson 2001)[173].

The listener is moved by the music from cognitive consciousness to phenomenal consciousness or even so called alternate states of consciousness such as trance and hypnoses. Such states heighten the suggestivity.

In the dyadic perspective the "living horizon of the now" (Husserl 1991: 16) contains the other as the frame of reference. Time is a shared dimension.

> …when people move synchronously or in temporal coordination, they are participating in an aspect of each others experience. They are partially living from the others centre. (Stern 2004: 81)

Emotional expression

"The emotions perceived to be expressed in musical stimuli are capable of evoking corresponding affective reactions in the listener" (Bruner II 1990)[174]. This simply states that music expresses the emotions we would expect it to express. If it is lively, it expresses liveliness and so on. But this argument is circular since it is lively because that is how we perceive it to be. We need a deeper understanding.

The mere fact that emotions can be expressed by music is a wonder. Music can express emotions just as faces, bodies and prosody

---

[173] Gabrielson, A. (2001) Emotions in strong experiences with music. *Music and Emotion – theory and research* (Eds. P.N.Juslin & J.A. Sloboda). Oxford University Press, Oxford.

[174] Bruner II, G.C. (1990) Music mood and marketing. *Journal of Marketing*. Vol. 54. No 4. 94–104

can. The whole register of emotions can be expressed. For this reason the key to understand music and emotion must be the understanding of emotional expression.

I am going to discuss vocal expression and body movement, but leave out facial expression. This is not because facial expression is unrelated to singing. I believe it is. But as yet there is no research on this.

Singing is obviously more basic than performing instrumental music, maybe with the exception of drumming. While playing an instrument usually involves declarative learning, singing is entirely procedural. Practically anybody can imitate a melody. We do not have to consciously learn how to move the body to produce a certain tone or sound. We can step up to the piano, hit a random key and then sing the tone, without testing. There is an automated connection between a heard tone and the motor production of this tone. A choir leader can give the voices tone and the whole choir starts singing at commando. Almost anyone can imagine a tone and then sing it. We can impersonate voices and accents. The body knows how, but we are not aware of this knowledge.

The most basic intention in vocal expression is to connect (Panksepp 1998)[175]. It is emotional before there is symbolical content. And when there are words, there is still always an emotional underpinning – the prosody. Is the vocal expression of emotion a common source for music and speech (emotional prosody)? This idea is far from new. As pointed out in the introduction Rousseau exercised this idea in *Essay on the origin of languages* 1761 (Rousseau 1986). The English philosopher Herbert Spencer stated 150 years ago in an article called *The origins and function of music* that there is a "general law that the feeling is a stimulus to muscular action" (Spencer 1857: 400)[176]. "Spencer's law" is still influential. There are several examples of involuntary connections between emotion and voice such as laughter, crying, sobbing, and the terrible experience of losing one's voice when in presence of authorities. How does this affect prosody

[175] Panksepp, J. (1998) *Affective Neuroscience – the foundations of human and animal emotions*. Oxford University Press. New York

[176] Spencer, H. (1857) The origin and function of music. *Fraser's Magazine*, 56. 367–408

and singing? Do we perceive sad music as sad because of the resonance with our own crying?

Laughter and crying is spontaneous. So is prosody. The emotion just comes out as behaviour. For the most part we are not even aware of it. And when we are, we have trouble hiding it. The expression reveals our state.

The topic of vocal expression in speech and music is a field of research. A review of recent literature is presented by Juslin & Laukka (2003)[177]. Laukka concludes in his dissertation (2004: 55)[178]:

1. Vocal emotions are understood cross-culturally (although there is an "in-group advantage")
2. Distinct patterns of voice cues correspond to *discrete emotions* (basic emotions such as anger, fear, disgust, happiness, sadness and tenderness).
3. The idea of an evolutionary perspective on vocal expression is supported.

Since voice cues are universally understood, there is no need to account for them here. They sound as we expect them to sound. It is enough to give an example. Sadness is, according to Laukka, expressed by

> [l]ow mean Fo, little Fo variability, low maximum Fo, falling Fo contours, little jitter, low voice intensity, little voice intensity variability, little high-frequency energy, low mean F1, wide bandwidth of F1, slackened articulation, rounded glottal waveform, slow speech rate, much pauses, and microstructural irregularity. (Laukka 2004: 56)

This simply states that sadness is expressed when the melody of speech is held in a low register, is not moving much and when it does, it is falling in small steps. The tempo of the sad voice is slow with pauses. The timbre is rounded and soft.

---

[177] Juslin, P.N. & Laukka, P. (2003) Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological bulletin.* Vol. 129. No 5. 770–814

[178] Laukka, P. (2004) *Vocal Expression of Emotion – discrete-emotions and dimensional accounts.* Acta Universitatis Upsaliensis, Uppsala

Intuitively we already know this. Furthermore the description fits with the expression of sad music (Gabrielsson & Lindström 2001). Chromatic (small second interval) falling movement is particularly characteristic of sad music. In fact all Laukka's findings about emotion in prosody are paralleled in the Gabrielsson & Lindström examination of emotional expression in music.

It does not take much fantasy to imagine the survival value of vocal expression of emotion and intention. I add intention here, because I believe that the mere intention to connect is vital. The urge of wanting something is typically expressed by inarticulate vocal expression (as any parent who has tried to pass the cash register in a supermarket with a five year old would subscribe to).

Recently the concept "the song system" or "the singing system" has appeared in neurological reports (Brown *et al.* 2004a[179]; Zarate & Zatorre 2005[180]). Does this imply that there in fact is a special song system in the sense that nature has designed a system for singing in human beings? We have, as with music, these two options:

1. "The song system" is an inherited system with direct survival value.

2. Singing is a target domain, a capacity depending on a composite of subsystems each with survival value developed in source domains (earlier developed capacities).

The first option is supported by the fact that singing is a universally distributed skill. It is, as stated, embodied procedural knowledge; and we have an inherited capacity to develop this skill. The second alternative entails that vocal emotional and intentional expression is the basic survival ability and that this ability together with other capacities has developed to the art form music.

Singing generally activates auditory cortices, primary motor cortices, the supplementary motor area (SMA), the insula, the cingulate

---

[179] Brown, S., Martinez, M., Hodges, D., Fox, P. & Parsons L. (2004a) The song system of the human brain. *Cognitive Brain Research*. Vol. 20. No 3. 363–375

[180] Zarate, J.M. & Zatorre, R.J. (2005) Neural substrates governing audiovocal integration for vocal pitch regulation in singing. *Annals of the New York Academy of Sciences*. Vol. 1060. *The Neurosciences and Music II: From Perception to Performance*. (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 404–408

sulcus, the thalamus, and the cerebellum (Perry *et al.* 1999[181]; Brown *et al.* 2004a; Zarate & Zatorre 2005). Additionally the inferior frontal gyrus is activated by singing (Riecker *et al.* 2000)[182]. This area is critical for pitch (Hyde *et al.* 2006) [183].

The insula and the cingulate cortex are not only activated by singing but also by music perception generally, (Koelsch *et al.* 2005; Tillmann *et al.* 2003a) and by pleasurable music specifically (Blood & Zatorre 2001[184]; Griffiths *et al.* 2004[185]; Brown *et al.* 2004b)[186]. These two brain areas emerge consistently in the reports of music production and perception. What do we know about them? Let me present some findings.

The insula and the cingulate cortex are generally known to be involved in the monitoring and control of bodily processes such as breathing. Together they constitute an area for perception and regulation of various bodily states including sexual feelings, appetite, taste, and the anticipation of potential gain (Davidson 2005)[187].

The insula is a large area situated behind the lateral sulcus. It connects the neocortex, the thalamus and the amygdala together. The general function is believed to be the production of emotionally relevant context for sensory experience. The frontal part called frontal operculum, also known as Broca's area, is important for sequenc-

---

[181] Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E. & Evans, A. (1999) Localization of cerebral activity during simple singing. *NeuroReport*. Vol. 10. No 18. 3979–3984

[182] Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G. & Heider, H. (2000) Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*. Vol. 11. No. 9. 1997–2000

[183] Hyde, K.L., Zatorre, R.J., Griffiths, T.J., Lerch, J.P. & Peretz, I. (2006) Morphometry of the amusic brain: a two-site study**.** *Brain*. Vol. 129. No 10. 2562–2570

[184] Blood, A. & 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*. Vol. 98. No 20. Sept. 25. 11818–11823

[185] Griffiths, T.D., Warren, J.D., Dean, J.L. & Howard D. (2004) "When the feeling's gone": a selective loss of musical emotion. *Journal of Neurology Neurosurgery and Psychiatry*, 75. 344–345

[186] Brown, S., Martinez, M. & Parsons, L. (2004b) Passive music listening spontaneously engages limbic and paralimbic systems**.** *Neuroreport*. Vol. 15. No 13. September 15. 2033–2037

[187] Davidson, R. (2005) http://www.eurekalert.org/pub_releases/2005–08/uow-sbs082605.php

ing speech and music (Lahav, Salzman & Schlaug 2007). This area receives projections from tongue nerves (Ogava, Ito & Nomura 1985)[188]. It has been pointed out as an area containing mirror neurons (Iacobini & Lenzi 2002). For this reason Iacobi & Lenzi have proposed that the insula is involved in empathic emotions.

The cingulate gyrus has been implicated in the mediation of social motivation – the desire to communicate (Panksepp 1998). This area is activated by music generally, and by rhythm specifically (Parsons 2003)[189] and as already pointed out by pleasant music. The anterior part is activated in singing. When electrically stimulated in this part subjects (squirrel monkeys) start involuntarily vocalizing (Muller-Preuss 1980)[190]. There is a feedback connection from the cingulate gyrus to the auditory cortex. Thus auditory cortex activity is altered by cingulate activity.

Both the insula and the cingulate cortex have been pointed out in an fMRI study of romantic love (Bartels & Zeki 2000)[191]. The stimuli were pictures of loved ones. As control, pictures of friends were shown. The romantic love foci overlap, according to the authors, with areas activated by drugs such as cocaine.

The insula and the cingulate gyrus overlap with the extrapyramidal motor system. For this reason Bartels & Zeki raise the question of whether the activity is related to motor planning or imagery associated with the loved person. These areas have not been reported before as activated in motor imagery tasks. An argument for the motor imagery stressed by the authors is that these areas, via the putamen, are part of the dopamine system for movement.

As can be seen from this brief survey the findings are quite disparate. But I think it is safe to state that the insula and the cingulate

---

[188] Ogava, H., Ito, S. & Nomura, T. (1985) Two distinct projection areas from tongue nerves in the frontal operculum of macaque monkeys as revealed with evoked potential mapping. *Neuroscientific Research*. Vol. 2. No. 6. August. 447–59

[189] Parsons, L.M. (2003) Exploring the functional neuroanatomy of music performance, perception and comprehension. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre) New York University Press. Chap. 17

[190] Muller-Preuss, P., Newman, J.D. & Jurgens, U. (1980) Anatomical and physiological evidence for a relationship between the "cingular" vocalization area and the auditory cortex in the squirrel monkey. *Brain Research*. Vol. 202. No 2. 307–315

[191] Bartels, A. & Zeki, S. (2000) The neural basis of romantic love. *NeuroReport*. Vol. 11. No 17. 3829–3834

gyrus are both implicated in the registration of bodily states as well as in the activation of vocal expression.

In this context it is interesting to recall that Panksepp emphasized that the cingulate cortex is activated by the urge to communicate. This implies what most singers and listeners feel: Singing is activated by the urge to express oneself socially – to communicate intentions and emotions.

I conclude: The insula and the cingulate cortex transform body states to emotions and emotions to vocalizations in the singer – and transform heard vocalizations to emotions in the listener. We know that these two areas are activated in the registration of

1. The own body (through feed back systems)
2. The other's body (through the mirror system)
3. Faces (loved ones)
4. Music

The connection between the neural registration of loved faces and vocalization in the same area is intriguing. I would say that the majority of all songs are about love. And that this is the best selling theme: love in every respect, but perhaps above all lost love. The imagined loved other.

This neurological exposé hints a possible understanding of Spencer's law. There is something in common not only in music and speech. Music is, as is prosody, after all heard movement. Let me remind the reader that the insula connects pre-motor areas with the limbic system. To some extent there is a common factor that effects vocal expression, body movement and music the same way – a general emotional effect on the body which propels movement patterns manifested in movements, prosody and music.

Emotional contagion

There is a difference between listening to somebody telling us she is happy and being affected by her smile. Sadly, we can be quite indifferent when somebody tells us about her happiness. This tells us that the information that the other is happy can be communicated on two different levels: As a semantic cognitive understanding emanating from the fact that someone is telling us about her state, or as so called *emotional contagion* defined by Hatfield *et al.* (1994: 5) as "the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures and movements with those of another person and, consequently, to converge emotionally".

Applied to music the distinction means that either we can explicitly draw a conclusion about the emotion: "This must be sad because it is in a minor key and the tempo is slow". This is a cognitive understanding. The mode minor and the tempo signify sadness. If we on the other hand just feel the sadness, without considering why, it is emotional contagion.

In the case of egocentric emotions we used the behaviourist model proposed by Edmund Rolls. Emotions according to this model are the result of rewards or punishment. In the dyadic case the emotion is not caused by a reinforcer. It is caused by another person contaminating us with her emotion. It is an emotion communicated. For this reason the behaviourist theory is not apt to describe dyadic emotions.

The emotional contagion theory is supported by evidence that we have special-purpose brain modules, which automatically and subcortically process acoustic stimuli, resulting in immediate emotional response (Juslin & Laukka 2004). The authors present the following list of evidence:

1. Modules are domain specific: There is brain dissociation between judgements of musical emotion and musical structure. (Peretz, Gagnon & Bouchard 1998)[192]. This implies that the emotion does not emanate from judgement of the structure.

---

[192] Peretz, I., Gagnon, L. & Bouchard, B. (1998) Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*. Vol. 68. No. 2. 111–141

2. Modules are quick: Judgements of musical emotions are quicker than cognitive judgements. (Peretz, Gagnon & Bouchard 1998).

3. Modules are innately specific: The ability to perceive emotions develops early. (Cunningham & Sterling 1988)[193]

4. Modules are autonomous: The processing in the perception of emotional expression is primarily implicit. (Niedenthal & Showers 1991)[194]

5. Modules are hardwired: It is impossible to relearn how to associate expressive forms with emotions. (Clynes 1977: 45)[195]

6. Modules are automatic: Emotional induction through music is possible even if listeners do not attend to the music. (Västfjäll 2002)[196]

Let me exemplify emotional contagion: If a person is thirsty and drinks – we need to drink, if she vomits – it makes us nauseous, if she urinates – we want to urinate, if she yawns – we have this sudden urge to yawn, if she cries – it brings tears to our eyes, if she laughs – she makes us laugh, if she gets angry – we might have a slight tendency to get angry too, if she sings a song – we tend to fall in, if she loves us and gives us a hug – it is easy to love her back, and if she is sexually aroused – we feel the urge.

Charles Darwin listed six categories of communicable emotions: disgust, fear, anger, sadness, tenderness and happiness (Darwin

---

[193] Cunningham, J.G. & Sterling, R.S. (1988) Developmental changes in the understanding of affective meaning in music. *Motivation and emotion.* Vol. 12. No.4. 399–413

[194] Niedenthal, P.M. & Showers, C. (1998) The perdeption and processing of affective information and its influences on social judgement. In *Emotion and social Judgements* (Ed. J.P. Forgas) Pergamon Press, Oxford, England. 125–153

[195] Clynes, D.M. (1977) *Sentics: The Touch of Emotion.* Anchor Press/Doubleday, New York

[196] Västfjäll, D. (2002) A review of the musical mood induction procedure. *Musicae Scientiae.* Special Issue. 173–211

1872)[197]. There are as Darwin proposed vital functions in emotional communication between conspecifics. If the other shows

1. disgust, you fear eating the same food.
2. fear of something, you get scared or you take advantage.
3. anger, you prepare to fight or to flee.
4. sadness, you tend to be sad or you might experience love and concern.
5. tenderness, you love her back.
6. happiness, you feel that everything is OK.

Since these emotions can be communicated it is enough that one individual discovers the danger; it is enough that one individual finds out that this can be eaten but not this etc. This is the evolutionary aspect of the benefits of communication of emotions. These emotions are expressed in a way that makes them communicative. As can be seen from the examples complementary emotions are sometimes aroused. The concept of emotional contagion is for this reason not optimal to describe direct and unmediated communication of emotion.

Empathy

Empathy (from the Greek εμπάθεια, "to suffer with") is commonly used in the sense of the direct experience of the emotion of the other – to "put oneself into another's shoes". As the reader may recall, imagination was described as the integration of a perspective into the egocentric perspective. The formulation "put oneself into another's shoes" catches this. Empathy thus depends on your ability to imagine that you are the other.

The Germans use the related word Einfühlung in the psychology of aesthetic experience, to denote the relationship between an art-

[197] Darwin, C. (1872) *The Expression of the Emotions in Man and Animals*. J. Murray, London.

work and the observer, who imaginatively projects herself into the contemplated object (Gallese 2003a)[198]. Theodor Lipps extended the concept of Einfühlung to the domain of intersubjectivity, which he characterized in terms of inner imitation of the perceived movements of others. When I am watching a tightrope dancer, Lipps noted…" ich fühle mich so in ihm" (Lipps 1903)[199].

The ability to take the perspective of the other is crucial for empathy. The importance of imitation to empathy is stressed by the notion that expression, and recognition of expression and gestures are impaired in individuals with autism. And, even more, by the observation that autistic infants lack in coordination of activity with their caregiver. Bråten observed that when a mother holds up her palms, children usually put their palms so that the palms touch. But children diagnosed as autistic tend to put their hands with their palms facing themselves (Bråten 1999)[200]. They imitate from an egocentric perspective! This indicates that the condition for empathy as the ability to take the other's perspective could be taken literally – not only metaphorically.

In 2002 Stephanie Preston and Francis de Waal published an article on empathy introducing the *PAM – model* (perception-action model) for empathy. The authors are influenced by the findings of mirror neurons. The article treats empathy as a communication of an emotion from one person to the other so that the emotion evoked in the subject is the same as the emotion of the other. The authors argue

> [that] attended perception of the object's state automatically activates the subject's representations of the state, situation, and object, and that activation of these representations automatically primes or generates the associated and somatic responses, unless inhibited. (Preston & de Waal 2002: 4)

---

198 Gallese, V. (2003a) The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology.* Vol. 36. No 4. 171–180

199 Lipps, T. (1903) Einfühlung, innere Nachahmung und Organempfindung. *Archiv für gesamte Psychologie.* 1. 465–519

200 Bråten, S. (1999) *Intersubjective Communication and Emotion in Early Ontogeny.* Cambridge University Press

The authors present a list of catalysts facilitating empathy:

1. Similarity
2. Familiarity
3. Cue salience

This article caused a storm. Psychologists argued that empathy is caused by cognitive awareness of the other's situation and thus that empathy is more than an implicit feeling. Researchers in the field of mirror neurons on the other hand were generally more in agreement with the article.

The emotion-movement connection

The communication of empathy thus depends on the automatic two-way connection between emotions and movements. *Proprioceptive feedback* (feedback informing the brain of the positions of body parts) is involved in this process. This can be exemplified: When we scratch a dog behind her ear, she starts kicking. And just by performing a facial expression mechanically, we get affected. This feedback from the manifestation of an emotion to the emotion itself is one of the keys to the understanding of the communication of emotions. This adds to an understanding how a perceived manifestation of an emotion in another person can evoke that same emotion in the beholder – without any intermediaries in the form of conscious representations. The model suggested by Preston & de Waal suggests that imitation explains how we are emotionally contaminated by the other's actions. The action of the other triggers a spontaneous tendency to imitate, and this tendency makes us feel the emotion, they claim.

The connection movement-emotion has been suggested before. William James presented the idea that emotions emanate from movements in 1884 in an influential paper called *What is emotion?* (James 1884)[201].

---

[201] James, W. (1884) What is emotion? *Mind.* 9. 188–205

Using the technique of Johansson (1973)[202], with point-light displays placed on people in dark environments in order to capture pure movements, Cutting & Kozlowski (1997)[203] found that we are not only capable of recognizing the identity of a walking person, but also the emotion of this person from pure movement observation (cheerful or depressed and sad).

As mentioned above, developmental psychologists have taken interest in imitation. Intuitive *motherese* (the way mothers address their babies) is universal. There is, says Daniel Stern, a focus in this conversation on the time course of the feeling intensity, the *temporal feeling shape* (Stern 1995)[204]. The temporal feeling shape of the observer will correspond to that of the observed, resulting in *affective attunement* – shared affective states. Daniel Stern calls this emotion the *vitality affect*, captured by kinetic terms: crescendo, decrescendo, fading, exploding, bursting, elongated, pulsing, wavering, effortful and easy (Stern 1993)[205]. As can be seen these terms partly correspond to the Italian terms for musical articulations.

The vitality affect adds a dynamic, cross-modal aspect to the understanding of empathy. It is, as is empathy, a dyadic emotion. It requires shared "almost identical capacities for temporal discrimination" between the participants. Stern shows how emotions are cross modally evoked, shared and regulated. The inner subjective qualities of feeling are displayed as form, intensity and timing, observable and directly perceived as a physical expression of emotions. For instance, the child smiles to her mother, the mother intuitively answers y-e-a-h in a smoothly arched pitch form showing the rise and fall of the smile of the baby, meaning – "Oh! This is pretty much how you must have felt, and I am showing you I know by matching the temporal contour of my activity to yours" (Stern 1993).

---

202 Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*. 14. 201–211

203 Cutting, J.E. & Kozlowski, L.T. (1977) Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*. 9. 353–356

204 Stern, D. (1995) *The Motherhood Constellation*. Basic Books, New York

205 Stern, D. (1993) The role of feelings for an interpersonal self. *The Perceived Self* (Ed. U. Neisser). Cambridge University Press

This "vitality affect" is empathetic. Stern's contribution is that he makes this emotion dynamic and multimodal (the contour of the smile is imitated with the contour of the voice).

Since the mirror neuron system is not only causing imitative action but also complementary action, we may suspect that this action leads to complementary emotions by means of proprioceptive feedback. But this intriguing possibility has up to now not been observed or examined.

* * *

How does the Preston & de Waals PAM theory fit with musical emotions? Research in music psychology shows that listeners reliably discern emotional categories from musical features such as:

1. amplitude – round/sharp
2. articulation – staccato/legato
3. harmony – simple/complex and consonant/dissonant
4. volume
5. dynamics – range in volume
6. melody – pitch, range, direction and intervallic steps
7. mode – major/minor
8. rhythm – regular/irregular, smooth/rough, complex/varied/firm, fluent
9. tempo – fast/medium/slow
10. timbre – soft/sharp
11. tonality – tonal/atonal
12. musical form

The richness of variables studied (multivariate analysis technique) gives us a hint of how many variables we "weigh together" in our perception of music without being conscious about it. Still these are very rough measures. By varying these variables different kinds of emotions can be perceived since every emotion has its own profile.

But, when we take a closer look at these musical parameters, we find that most of them could be analysed as movements. Intervallic leaps are the size of the movement. Legato is a description of the smoothness of the motion. Volume is, as pitch, and speed related to intensity of the motion. Dynamic change is the vitality contour of motion, and so on. The same could be said about prosody and of course body language. From this point of view Spencer's law seems to be correct.

We can now see the importance of immediate imitation – the synchronization of movements. In fact we have to share the present moment to be able to apprehend the change in intensity. We perceive each other directly. Let us review Hatfield's definition! Emotional contagion is described as "the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures and movements with those of another person and, consequently, to converge emotionally". Read literally, synchronizing leads to emotional contagion. Although radical, the statement fits with the feedback connections between emotions and their manifestations, described above. The importance of the time factor is verified by Stern's video experiments of mothers presented on screen to their babies. When a short time displacement between the baby's utterance and the mother's reaction was made, the baby lost interest. Even if there is much more to explain in this process, we cannot ignore that in a broad field of research immediate imitation (synchronization, entrainment) is a condition for communication of emotions. It is the common factor.

Neurological support for the PAM model?

Koelsch (*et al.* 2006)[206] detected "a motor-related circuitry for pleasurable music comprising the ventral striatum, the insular cortex, the Rolandic, and possibly the frontal operculum". This circuit served the formation of pre-motor representations for vocal sound production during the perception of the pleasant musical excerpts. This indicates a specific perception-action network for music listening and vocal production.

This interpretation is supported by a study reporting activation of the Rolandic operculum during the processing of intonation contour while listening to spoken sentences (Meyer *et al.* 2004)[207], suggesting that pre-motor programs for the production of prosodic signals (such as speech melody) are already formed when we are perceiving prosodic signals produced by other individuals.

According to Koelsch the activation of the Rolandic operculum was observed during the perception of the pleasant, but not of the unpleasant musical information, "presumably due to the appetitive quality of the pleasant excerpts (the unpleasant excerpts rather evoked avoidance behaviour)". At first sight this seems to be an interesting thought. It underlines the emotional (empathic) aspect of implicit imitation. Empathy is embodiment. But this argument comes dangerously close to being circular. If imitation causes pleasure, pleasure cannot cause imitation. If the appetitive quality is caused by imitation, it cannot be the reason for it. I would suggest another interpretation: If we imitate the music, we like it. What then decides if we imitate or not? I will come back to this in the concluding chapter.

This study supports the idea that music emotions imply mirror neurons, the insula and the limbic system, but it is noteworthy that it is not supportive of the PAM model. The emotion described by Koelsch is not empathy in the sense that the same emotion is

---

[206] Koelsch, S., Fritz, T., v. Cramon, Y., Müller, K., & Friederici, A.D. (2006) Investigating emotion with music: An fMRI study. *Human Brain Mapping*. Vol. 27. No 3. 239–250

[207] Meyer, M., Steinhauer, K., Alter, K., Friederici, A.D. & von Cramon, D.Y. (2004) Brain activity varies with modulation of dynamic pitch variances in sentence melody. *Brain and Language*. Vol. 89. No. 2. 277–289

evoked in the beholder as in the other. Rather he describes a general positive emotion. If so mirror neurons could not explain why different kinds of music elicit different emotions. But it could explain the general good feeling elicited by music even when the music is recognized as sad.

Conclusion

In Chapter 1 I argued that perceptual learning is based on action and thus that it implies a movement pattern. The knowledge of sounding music is procedural. We called this pattern a sensory-motor schema. In this chapter we have discussed a mechanism implied in action imitation and completion – the mirror neuron system. This is a pre-motor-insula-limbic circuit. There is thus neurological support for the theory that movement and emotion are reciprocally connected. In the dyadic perspective, the listener imitates the emotional expression of the other (e.g. the musician) manifested in the music. This imitated movement causes the emotion.

The fMRI study of Koelsch supports the theory to some extent, but it stresses that mirror neuron activity causes a general positive emotion and that there is no mirror neuron activity if the music is not liked. If so mirror neuron could not explain the variety of emotions in music listening and the PAM theory could not be applied to music. The research is however in a virginal state. I foresee that we will have clarifying evidence within a few years.

Discussion

Mirror neuron theory explains implicit imitation and completion. This is a prerequisite for communication and learning. It makes us sense the intentions, the goals and the emotions of the other directly and implicitly.

These points could preferably be combined. The emotional bond is vital for learning and it is vital to have an emotional understanding of the situation to see the purpose of action. Thus it has been argued that covert imitation plays a role in our ability to understand other people's actions, so called action understanding (Gergely & Csibra 2003)[208], which involves the intentions and predicted goals behind actions (Buccino *et al.* 2004)[209]. The most full-blown theory suggests that the observer, in her reproduction of the other's movements, reproduces the sensory consequences of those actions (Hesslow 2002)[210]. Wilson and Knoblich (2005) suggest that we use our own bodies to track the other person's actions in real time and that this produces a sense of top-down understanding (I suggest imagery) filling our blanks in the sensory information and suggesting predictions. It actually means that we transpose what we see in others to action tendencies in our own bodies and that this triggers imagination of intentions and outcomes. This produces top-down expectations and predictions of the unfolding action.

In *Vom Musikalisch Schönen* Eduard Hanslick stated:

> Was kann also die Musik von den Gefühlen darstellen, wenn nicht deren Inhalt? Nur das Dynamische derselben. Sie vermag die Bewegung eines physischen Vorganges nach den Momenten: schnell, langsam, stark, schwach, steigend, fallend nachzubilden. Bewegung ist aber nur eine Eigenschaft, ein Moment des Gefühls, nicht dieses selbst (Hanslick 1986: 26).

As we can see Hanslick's concept "*Moment des Gefühls*" is close to Stern's vitality affect. Hanslick did not and could not know about the connection between the perception of "*den Momenten: schnell, langsam, stark, schwach, steigend, fallend nachzubilden*" and emotion. I suggest that one dimension in music is the vitality affect and that this affect is picturing the composer's and the artist's sensorimotor schema. This can be captured in a music score or a phonogram and brought to life later and act on the sensory motor schema of the listener. To be contaminated by music therefore does not necessarily

---

208 Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences.* Vol. 7. No 7. 287–292

209 Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language.* Vol. 89. No. 2. 370–376

210 Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences.* Vol. 6. No. 6. 242–247

demand the presence of *the other*, even if this helps as additional information and action understanding. When we listen to a record the artist is absent, so is the composer, still we are entrained by the flow.

As we have seen the perception process is to a large extent pre-attentive. This means that we do not have to be aware of the other to be emotionally affected. It is thus not critical that we perceive the other in the music consciously. In other words we do not have to think or experience the music as the expression of the other to be affected by it – the movement of the other affects us anyway. The emotion does not result from the appraisal of the other; it does not result from conscious judgements.

A lot of work remains to be done to confirm this theory. The evidence is, as we have already seen, not unambiguous. Let me provide one more example. In the 2007 as well as the 2005 investigation of auditory mirror neurons, presented above, Lahav and his colleagues found that the subjects did imitate the music implicitly, provided they had rehearsed the piece and had the procedural knowledge to play it.

Although this seems to confirm the hypothesis, it does in fact contradict it. Lahav's test shows that we have to learn how to play the music for the activation in the mirror system to occur. This is the problem. It means that the mirror system is not activated in most cases. It has been said that Rachmaninov's four piano concertos are very emotional, but how many in the audience knows how to play them?

On the other hand Zatorre (*et al.* 2007)[211] report that "even when listeners do not have explicit sound-movement associations, such as when passively listening to rhythms in a naive condition without foreknowledge about any motor task, they still show recruitment of pre-motor cortices and the SMA".

The last finding undoubtedly opens up for the possibility that we imitate unrehearsed pieces. We could imagine that even if we do not know how to play a song, we might know how to sing it and

[211] Zatorre, R.J., Chen, J.L. & Penhune, V.B. (2007) When the brain plays music: auditory– motor interactions in music perception and production. *Nature Reviews. Neuroscience.* Vol. 8. No 8. 547–558

beat the rhythm without rehearsal. ERP-experiments have shown that we can predict music if we are familiar with the type of structure. We will come back to this in the following chapter.

This calls for further investigations. Here are some questions that have to be answered:

1. Is the activation of the mirror system a condition for musical emotions? This demands a test where mirror system activation is correlated to physiological responses.
2. Can the mirror system be activated by songs that we know how to sing, dances that we know how to dance or symphonies that we know how to conduct?
3. Can the mirror system be activated by music that we do not know but that is easy to predict (pop music as compared with atonal music)?
4. How is imitation and prediction related? Are these two separate processes? If so can we feel the mismatch when we imitate some thing we did not predict?
5. How is the rather mechanical theory of entrainment related to imitation and prediction?

Let us weave the proposals made in this chapter together in one example! You are listening to Mozart's *Ein Weib ist das herrlichste Ding*[212]. Now let us put the composer in the frame of reference. The phenomenologist Alfred Schütz once wrote:

> Although separated by hundreds of years, the latter (the listener) participates with quasi-simultaneity in the former's (the composer's) stream of consciousness, by performing with him step by step the ongoing articulation of his musical thought. The beholder, thus, is united with the composer by a time dimension common to both. (Schütz 1964: 171)[213]

---

[212] Mozart, W.A. CD: Piano Variations. Vol. 3. Naxos

[213] Schütz, A. (1964) *Collected papers: studies in social theory.* Vol 2. (Ed. A. Brodersen). Martinus Nijhoff, Hague

As mentioned in Chapter 1, imagery is the only way we are able to grasp structure in music (Halpern 2001). Thus composers must work in real time when they create music. Imagination elicits implicit movements. In other words, the composer, having the procedural knowledge of music, starts a real time imagery of music. You entrain the music and thereby the real time implicit movement of Mozart as he wrote the piece in 1791. You follow his imagination of the unfolding music and you share his bodily sensations (latent movement tendencies). You embody him. You are, in this sense, Mozart. This is an experience that separates music from other art forms. It depends on the *temporal topography* of music (Kivy 1989).

And, this chapter argues, this is the reason why you momentarily share Mozart's emotion.

4

Navigating in the Soundscape

The Allocentric Perspective

There is a resemblance between following a path travelling through a landscape and following a melody as it meanders through the musical topography. The structure unfolds. In this respect the listener can be compared to a the wanderer on a path, the rat in a maze, the driver on the highway, or the player passing through a virtual landscape in a computer game.

Do we perceive music structure the same way we perceive other structures? For a structure to appear we need fixed landmarks in that structure. The perception of the topography in a landscape demands allocentric landmarks.

The chapter consists of two parts. In the first part the perception of music structure is discussed. In the second part I will point out the parallels between the perception of music structure and the perception of structure in the landscape - navigation. I will argue that we have allocentric representations of music structure in the brain.

Depiction of perspective

The allocentric perspective is taken from a landmark situated outside the body. This has consequences for our perception of topography and our ability to navigate.

Examples:

1. The location of the shop on my way to work can be perceived either as “on my left” or opposite the church. In the first case I take an egocentric perspective using my own body as a landmark and in the second case the perspective is allocentric since the landmark is allocated to the church.

2. When you walk in your home village you have a sense of direction. This is because you relate to some allocentric landmark such as the church, the nearest town, the cost etc. Your movement is defined in relation to this landmark.

3. When you turn, the egocentric perspective changes. What was on your right is now on your left and vice versa. But the allocentric perspective remains the same, because you are on the same side of the church however you may turn.

4. In the egocentric perspective the world is variant (changing) as you walk. This is the consequence of carrying the landmark with you. In the allocentric perspective the world is invariant (stationary), provided the landmark is stationary, and you are moving in the world. Direction and speed are related to the landmark. From an allocentric landmark a fixed structure, a *topography*, appears.

The music structure consists of heard events on a temporal dimension. We must be able to identify places in the music for the structure to appear and relate musical events to these places. I will discuss the possibilities that the tonal centres and the sense of temporality in music produce allocentric landmarks.

This discussion raises the question in what sense we can talk about place in sounding music. As stated in Chapter 3, spatial dimensions must not be seen. They can be heard. This was discussed as the perception of the auditory scene. We can point out directions and tell the distance to the different instruments in an orchestra. This is however not what we are discussing in this perspective. Our concern here is the soundtrack as presented from one loudspeaker. It is the temporal stream of music, not a composite of sounds distributed in space. Even so, we perceive place. In this respect the sound-track is to the listener what the scent-track is to a dog. The question about the perception of place in music will be illuminated in this chapter.

The perception of music structure

All music is structured along some organizing principle. If not, we would just experience chaotic noise. The organizing principle varies with music culture and styles.

How is music structure perceived? On a basic level structure is perceived as reactions to the unexpected. The expected is seldom noted. If a tooth is missing in a smile, this is the tooth we will notice. The missing tooth indicates that we were mistaken about the structure and that we have to reconcile. Thus we do not hear every instrument in an orchestra. But if one of them is playing out of tune, it will be heard. The fact that mismatches direct our attention makes us remember mismatches easier than matches. The matches are so called redundant information that does not draw our attention. Since mismatches are attended to, there will be an effect on the previously presented feature imprinting (perceptual memory).

Structure is thus perceived indirectly. The mere fact that we react to the deviance (the inappropriate chord, the false tone, the misplaced drumbeat) shows that we have an expectancy of structure and thus some kind of representation (memory) of music structure.

In order to perceive structure, the allocentric perspective must be integrated with the egocentric perspective. The quasi-spatial dimensions are lived as events in the narrow window of time called the present moment. The time window limits the perceived structure to a few seconds. Do we feel longer structures? If we did not, music form would not affect us. The sonata form, to take an example, can be discussed theoretically, but is it felt by the listener? Would it matter to a listener not educated in music form? This has to my knowledge not been tested.

The question is: How far ahead can we anticipate? My experience tells me that musicians feel at least 16 bars. In order to play by ear in a group, musicians must know where they are. If they are experienced, they do not have to count bars to know this. I imagine that conductors need to have an intuitive "grip" of the music - the breathing of the entire structure. Gunnar Eriksson, conductor of Gothenburg Chamber Choir, aims for the subdominant. When he arrives there with the choir the rest of the phrase will unfold by itself. And this aim for the subdominant is what he tries to communicate to the choir (Eriksson 2008)[214]. Likewise, in order to compose by ear or improvise, musicians must be able to feel the structure far ahead. It has happened that I have woken up with a 16 bar melody clear and ready. The question is however tricky since, as I intend to show, we perceive structure on a preconscious level. It is thus hard not to confuse cognitive knowledge of the piece with perceptual anticipation. The first kind is declarative and the second is probably procedural. If we are to feel these longer structures in the present moment then unitizing mechanisms are needed. I will discuss five main organizing principles: gestalt**,** perceptual learning, sequencing, chunking and music rules (of discrete pitch and time, tonality and periodicity).

Gestalt

Leonard Meyer was the first musicologist to systematically explore Gestalts in music. In his classical book *Emotion and meaning and music* (Meyer 1956) Meyer argued that the listener would react emotionally to fulfilments and violations of expectancies emanating from the sense of Gestalt.

It has been demonstrated that the Gestalt principles act as organizing factors in music perception (Bregman 1990). However, since Gestalt laws are universal, we would perceive unfamiliar music just as well as we perceive familiar music, if this was the only structuring principle. This is not the case at all. We have to look for learned effects.

[214] Eriksson, G. (2008) Personal information. 12/1

Perceptual learning

From exposure to melodies, we learn how a heard melody is likely to develop. We have, according to David Huron, learnt to expect small melodic leaps rather than big leaps in Western music. Large intervals are more likely to ascend than small intervals; whereas small intervals are expected in descending melodies. If the melody is falling in small steps, we expect it to continue to fall. Large leaps are likely to be followed by a change in melody direction. A phrase is likely to have the shape of a bow. The end of a phrase is expected to fall (Huron 2006).

This type of learning exemplifies a category of perceptual learning called "topological imprinting" (c.f. Chapter 2).

We should separate perceptual learning within one piece of music from generalization between pieces. Learning one piece is to learn a sequence. When the knowledge from one piece is applied to another piece, musical rules (sense of style) are formed.

Sequencing

How can we explain that jazz and rock musicians can play entire concerts without music scores? The information processed is so immense that we can rule out that the whole structure is remembered. The speed of the soloist is too fast for conscious decision taking. We must assume automated chains of sequences and a set of cues to start these chains. There is only place for conscious considerations on a large scale. This is like remembering a telephone number. The number is at the same time individual digits and an entity. When we hear the first digits, we know the rest. Sequences are thus learned through the uniting *chunking* mechanism in perceptual learning (Smith 1999)[215].

Bob Snyder has applied the theory of chunking to music learning (Snyder 2000). Chunking make us perceive separate musical ele-

[215] Smith, G.J. (1999) Teaching a long sequence of behavior using whole task training, forward chaining, and backward chaining. *Perceptual and Motor Skills*. Vol. 89. No 3. 951–965

ments as parts of larger entities. These entities can be short themes, ostinatos, funky riffs, vamps of chords, or a repeated base line as in the passacaglia. Music is in fact a map of sequences. It is based on repetitions of limited material. Let us say that a vamp of four chords is four bars long and that this vamp is felt as one unit. Does this make it possible to perceive 16 bars of music?

Sequences are basic structures in music. A sequence is a definite structure. When we hear the first notes we know what to expect. However the sequences adjust to the harmonic context, but let us leave that for the moment.

Sequences vary along a "What" and a "When" dimension. In music the dimensions time and pitch are reflected in the musical score, which can be seen as a rough diagram where events are plotted in frequency-time space. There are however many parameters in music such as harmony, timbre, volume etc. This calls for a multidimensional understanding of music sequences. Let us simplify this by calling music sequences *a fixed repeated order of sound events along a temporal dimension.*

Time has three aspects: The order of events, the duration between events, and tempo – a pattern of intervals constituting the beat of time (c.f. Chapter 3). In speech the order is critical for understanding but the duration between events is not as important and not as precise as in music. Music sequences differ from speech sequences in this respect. If the tone values (duration) are changed, the identity of the song is changed, while such a manipulation would not affect a spoken sentence as much. This is reflected in the fact that musical scores have time specifications and written language has not.

The dimensions of a sequence can be studied one dimension at a time (e.g. the same tone in the ordinal dimension or an even beat in the temporal dimension). However, there is a strong transfer effect – one dimension supports the other (Ullén *et al.* 2005). Timing has a role in the perception of pitch. We are better at judging pitch if we know when the tone will appear (Huron 2006). The cerebellum is activated by pitch determinations (Parsons 2003). Since the cerebellum is involved in rhythm processing, this finding may support the transfer effect.

Musical rhythm is not just a division of time. It is a sequence – repeated sound events distributed over time. Drums and percussion instruments have different timbres and different volume. Drummers tune their drums to get the right pitch. Furthermore the drummer varies attack and place where the drum is hit. There is already in the musical rhythm a variety of sound events. Musical rhythm thus is multidimensional. In fact it is quite difficult to beat a drum without grouping the beats into subdivisions, marked by differences in volume. Even as we walk, we tend to feel one-two, one-two-three or one-two-three-four where the "One" is a little heavier.

We can keep short musical sequences in our working memory for a while. In order to remember an entire song a row of sequences have to be connected to each other and the whole thing must be stored in long term memory. Our ability to learn sequences has been studied in the serial reaction-time task (Clegg *et al.* 1998)[216]. In this test subjects are told to point out dots on a screen when they appear. If there is a repeated pattern of the appearance of the dots, the subjects become faster in their performance, even if they are not aware of the sequence (Shin & Ivry 2002).[217]

Sequences thus can be perceived and learned, although the subjects are not aware of the existence of the sequence. The confirmations of expectances are called *matches* (between mapped representations and the perceived environment). When the information does not fit the expected structure, we have a *mismatch*. The mismatch triggers a neurological reaction. Such reactions are automatically retrieved. But the mismatch as such often directs our attention to the event. It is an indication that unforeseen information is at hand.

We can see this mechanism in dogs and horses. They react strongly when something unexpected appears along a familiar path. They do not react to the object as such. It could be something harmless as a bike thrown in a ditch. It is the mere fact that this thing was not here before that is alarming. In this respect we do not differ

---

216 Clegg, B.A., DiGirolamo, G.J. & Keele, S.W. (1998) Sequence learning. *Trends in Cognitive Science.* Vol. 2. No 8. 275–281

217 Shin, J.C. & Ivry, R.B. (2002) Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition.* Vol. 28. No 3. 445–457

from the animals on a pre-attentive level. The unexpected is always somewhat alarming.

We can pass a city driving a car, while we listen to the radio or talk to a passenger. We take hundreds of decisions automatically. The flow of the traffic is unattended. But when the unexpected happens, we are immediately directed to the situation at hand. The surprise is arousing. This changes the perspective. What was previously in the background is now figure.

Pre-attentive surprise reactions can be measured. The technique is called ERP (c.f. Chapter 1). ERP is a change in the EEG pattern. Originally this was done with so called *odd balls* (i.e. an unexpected red ball in a sequence of blue balls).

This technique has been used to test reactions to deviances from anticipated musical sequences (Tervaniemi 2003). The findings show that we do react. ERPs are *pre-attentive*. Subjects asked not to pay attention to music sequences, but to read an interesting book by their own choice, showed significant ERP reactions on unexpected tones (Brattico *et al.* 2001)[218]. This indicates that we are able to follow sequences without paying attention, just as we can follow a path absentmindedly. These experiments show three interesting things:

1. Sequences can be learned pre-attentively.
2. We have (for a while) a representation of this sequence.
3. The representation creates expectations (attended or unattended).

The processing of auditory sequences has been studied by neurologists. Except for the obvious fact that the auditory cortex is involved, we can discern a sequence circuit comprising:

1. Motor, supplementary, sensorimotor and pre-motor areas and cerebellum (Jeannerod & Frak 1999[219]; Dhamala *et al.*

---

[218] Brattico, E. Näätänen, R. & Tervaniemi, M. (2002) Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception*. Winter. Vol. 19. No 2. 199–222

[219] Jeannerod, M. & Frak, V. (1999) Mental imaging of motor activity in humans. *Current Opinion in Neurobiology.* Vol. 9. No 6. 735–739

2003[220]; Chen *et al.* 2005)[221]. The supplementary and the pre-supplementary motor areas are involved in chunking (Kennedy *et al.* 2001)[222].

2. Basal ganglia (Janata *et al.* 2002)[223]. This area is critical for timing and attention. It has been shown that attention is higher when tones are expected. Thus we discriminate pitch better when the tone event occurs when anticipated (Jones *et al.* 2002)[224]. Striatum (in the basal ganglia) has been pointed out as crucial for sequential movements (Quartz & Sejnowski 2002). Comment: This is in accordance with the transfer effect discussed above.

3. Prefrontal cortex involvement in music sequencing anticipatory tasks has been confirmed by an fMRI study. This part is generally known to be critical to foresight. Specifically parts of this region is activated in music imagery tasks (the right inferior and middle frontal gyri (BAs 45, 46, and 47) (Halpern & Zatorre 1999)[225].

In sum, neurological evidence indicates that the processing of sequences involves brain areas associated with anticipation, attention, movement, timing and ordinal mapping. There is a transfer effect

---

[220] Dhamala, M., Pagnoni, G., Wiesenfeld, K., Zink, C., Martin, M. & Burns, G. (2003) Neural correlates of the complexity of rhythmic finger tapping. *NeuroImage.* Vol. 20. No 2. 918–926

[221] Chen, J., Penhune, V. & Zatorre, R.J. (2005) Tapping in synchrony to auditory rhythms. *Annals of the New York Academy of Sciences.* Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 400–403

[222] Kennerley, S.W., Sakai, K. & Rushworth, M.F. (2004) Organization of action sequences and the role of the pre-SMA. *Journal of Neurophysiology.* Vol. 91. No.2. 978–993

[223] Janata, P., Tillmann, B. & Bharucha, J.J. (2002) Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cognitive, Affective & Behavioral Neuroscience.* Vol. 2. No 2. 121–140

[224] Jones, M.R., Moynihan, N., MacKenzie, N. & Puente, J. (2002) Temporal aspects of stimulus driven attending in dynamic arrays. *Psychological Science.* Vol. 13. No 4. 313–319

[225] Halpern, A.R. & Zatorre, R.J. (1999) When that tune runs through your head: a PET investigation of auditory imagery of familiar melodies. *Cerebral Cortex.* Vol. 9. No. 7 October. 697–704

between pitch and rhythm. The sequence is learned as an entity in the sense that when the first units are perceived the sequence unfolds.

Rules

As already stated all music is structured according to some kind of organizing principle. These principles can be seen as rules.

The mechanisms treated so far (Gestalt, perceptual learning, chunking, and sequencing) contribute to the formation of music rules. Rules are generalizations. Rules create, as do sequences, expectancies. But the difference is that a sequence concerns the single case whereas a rule concerns all cases. In a sequence we know that, if *this*, then *that*, whereas a rule states that if *this*, then *that* or *that* or *that,* but not *that*. Sequences do not permit variation. Rules do. They permit everything that is not forbidden. A hand of cards is a sequence. Rules come into force when we decide to play bridge. In traffic a sequence can be a description of how to get to the football stadium, whereas traffic rules imply that as long as you follow the rules you can go anywhere.

Rules are not just explicit, as traffic rules are. We have an implicit sense of rules. Rules help constitute the frames of reference by which we perceive the world, understand the world and expect continuances in the world. In the case of music this sense of rules is developed very early: "[b]abies nervous systems seem to be equipped with a capacity to sort out the different musical sounds reaching their ears in order to construct a "grammar", or a system of rules (Zatorre 2005)[226].

When a musician signals her intentions to the others in the group (e.g. this is the end of my solo; here comes the chorus) by playing an agreed figure, this is called *cuing*. To cue is to present the "*this"* in the "*if this, then that"* rule. This is an explicit rule, but there are auditory cues leading to anticipations on every level of consciousness.

226 Zatorre, R.J. (2005) Music, the food of neuroscience? *Nature.* Vol. 434. No 7031. 312–316

Examples:

1. The rule of discrete pitch and time
2. Genre and style derived rules
3. Tonality (or other hierarchical organization principles)
4. Periodicity – rules of repetition and variation
5. Rules of voicing
6. Rules of harmony
7. Rules of rhythm-harmony-melody connections
8. Rules of improvisation

The listed categories are not mutually excluding. The first four categories contain basic music rules and the following categories can be derived from these four. Voicing is, to take an example, dependent on discrete melody and time, on tonality, on style and on periodicity. David Huron (2001b)[227] listed thirteen general rules of voice leading that appear in pedagogical works.

To teach music is to a large extent to make the student aware of such rules. The music taught in music schools is a conceptualization of implicitly felt music rules. To some extent implicit rules have become explicit as a consequence of notation. The common factor of these rules is that they govern what is "allowed" and what is not "allowed" continuances in music structures. They are however not thought of as strict rules, since there is an aesthetical value in surprising continuances. Composing is to a large extent to relate to these rules. This is connected with creativity and taste. The composer or the improvising soloist may surprise, but preferably in style. The surprise is recognized by the connoisseurs. It has a strong emotional effect. In flamenco the unexpected move is commented by the afinados – *olé* – as recognition of *duende* (Spanish, literally

[227] Huron, D. (2001b) Tone and Voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*. Vol. 19, No. 1. 1–64

meaning goblin or elf – the presence of magic) (Martin 1973: 43)[228]. We will return to the emotional impact of surprises.

We shall now focus on the rule of discrete pitch and time, the rule of tonality and the rule of periodicity.

The rule of discrete time and pitch

Most music consists of sound events distributed over discrete scales and discrete rhythms. Björn Merker (2002)[229] argues that music could be considered to be a Humboldt system. A Humboldt system is a limited number of discrete elements that can be combined endlessly. Chemistry, genetics and language are examples of such systems.

Musical tones are however not unequivocally discrete entities in the same sense as the atoms in chemistry or the molecules in DNA. Discrete pitch and time in music is a rule, but a rule with exceptions. Music does not always have this discrete quality. In fact we often hear glissandos, blue notes, bended notes etc., not to mention false singing. A tone can occur anywhere. Likewise the tempo can be altered (rubato, accelerando, ritardando) and the durations in rhythm can be altered according to style. There are for example slight variations in the samba rhythm between the different samba schools in Rio de Janeiro. These variations, hard to detect for the tourists, are important emblems to the schools. Brazilian music shows many examples of so called malleable meter (Tupinambá de Ulhôa 1999)[230]. The main beat may be straight as a clock but the subdivisions are floating. It comprises polyrhythm and meter.

In written music such exceptions from discrete pitch and rhythm are often hidden. We just see how the two dimensions of pitch and time make up a two dimensional discrete universe. But the score is of course a simplified graph of sounding music.

---

[228] Martin, J. (1978) *El Arte Flamenco de la Guitarra*, United Music Publishers Ltd, London.

[229] Merker, B. (2002) Music: The missing Humboldt system. *Musicae Scientiae*. Vol. 6. Spring issue. 3–21

[230] Tupinambá de Ulhôa, M. (1999) Métrica Derramada: Prosódia musical na canção brasileira popular. *Brasiliana – Revista da Academia Brasileira de Música*. Vol. 2. 48–56

The discrete quality is in the listener's expectation. We anticipate discrete scale notes on the beats (or on subdivisions of the beat). It is the rule. When we hear a glissando, a rubato, or false singing, we feel the deviance from this rule.

This rule radically limits the anticipated continuation. Instead of an endless row of possible outcomes along a continuum of time and pitch, only scale tones related to the beat are expected.

Tonality rules

When we speak about tonality, we presume the harmonic sequences and melodies in major or minor scales leading to tonal centres (the tonic chord and the tonic note). It is a system of hierarchical pitch relationships.

Although the term tonality did not exist before the 19th century, tonal music developed out of the polyphone 17th century modal music in Europe. Tonality was born in the cadence – the technique to lead the voices to end a phrase – to bring it home.

The pitch relation in an interval can be consonant or dissonant. Consonant intervals are the octave, the perfect fifth and the third (the major third more so than the minor third). We could add the sixth and the forth, since they can be perceived as thirds and fifths (inversions). All the other possible intervals are considered dissonant. Music theorists argue that chords containing dissonant intervals "want" to be dissolved in consonant chords. People who have been exposed to tonal music automatically expect this to happen.

Experimental evidence shows that infants have early preferences in consonant intervals (Saffran 2003)[231]. Consonant intervals form consonant chords and dissonant intervals form dissonant chords. Dissonant chords are perceived as less pleasant (Blood *et al.* 1999). It has been shown that they activate the amygdala (unpleasant emotions) while consonant chords activate prefrontal cortex (possibly

---

[231] Saffran, J.R. (2003) Mechanisms of musical memory in infancy. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York

inhibiting amygdala anxiety). It is known that prefrontal cortex inhibits signals from amygdala (Quirk & Milad 2002)[232].

So we have a preference for consonance. Why is that? I discern three possibilities:

1. Exposition to music. We are more exposed to the first, the fifth and the third tone in the scale in western tonal music (Huron 2006). They are thus more expected whereas other tones are less expected. Surprises are generally on a basic level alarming and scary.

2. Exposition to sound. As stated in Chapter 2 the strongest tones in the overtone spectrum are the octave and the perfect fifth and the major third. It follows that these intervals are the most frequent in natural sounds. It is thus possible that our natural environment has created expectancies of consonant intervals.

3. The tones in consonant intervals have resembling overtone structures. The extreme case is the octave. Here the two tones have identical overtone structure. The octave is psychologically perceived as the same tone as the fundamental. The overtones in consonant intervals or chords do not clash as they do in dissonant chords. There is in other words a physical unevenness in dissonant chords.

The first alternative indicates that tonality causes preferences, whereas the other two alternatives indicate that preferences may have caused tonality.

The overtone structure is so vital for the perception of the fundamental that the fundamental is actually heard even if the sensory input only contains the overtone spectrum. I suggest that our ability to identify the rote in a chord is based on the same mechanism.

We have a sense of pitch relations and a sense of the rote in a chord. This is the basis of tonality. Tillmann (2005)[233] has shown that

---

[232] Quirk, G.J. & Milad, M.R. (2002) Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature*. Vol. 420. Nov. 7. 70–74

[233] Tillmann, B. (2005) Implicit investigations of tonal knowledge in non-musician listeners. *Annals of the New York Academy of Sciences*. Vol. 1060. *The Neurosciences*

non-musicians have implicit knowledge of tonal connections that are explicitly known by music theorists.

Krumhansl & Toiviainen (2003)[234] tested a group for the expected tone after a scale sequence. The authors found, as could be expected, that the first choice was the next tone in the scale, the second choice a chord tone, the third choice another scale tone, the forth choice tones outside the scale, and the fifth choice tones on fractions of intervals such as quarter tones. This leads to the possibility to construct a map of probability for different tones. Huron (2006) accounts for several behavioural studies, which support probability maps based on tonality. This leads to the conclusion that in a certain musical context some continuances are generally more expected than others.

In this sense tonality makes up a frame of reference, which gives music meaning. If we focus on the melody, the harmony will give the melody meaning. The implicit sense of music rules, makes us predict that the melody should be constructed of chord tones or colourings of the chords. This is the internal model of how it should sound. Violations of such rules affect the perception of the melody. This effect is evident in the advanced technique in jazz music called *playing outside* (Levine 1995)[235]. This means that an improviser leaves the harmonies and plays melody tones clearly excluded by the harmonic context. Then she returns to the harmonic context. When the melody is *out* it is at odds with the prediction produced by the internal model, whereas *in* is in concordance with this model. Playing *out* is perceived as a wild, free, forbidden and almost rebellious excursion and the return *in* is felt a return to law and order. Daddy is home. Playing outside the harmony often makes the player move sideways as if placing her outside the internal model. This movement indicates that the player has a bodily experience of when she is out and when she is in.

---

and Music II: From Perception to Performance (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo). 100–110

234 Krumhansl, C.L. & Toiviainen, P. (2003) Tonal Cognition. *The Cognitive Neuroscience of Music*. (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. 95–108

235 Levine, M. (1995) *The Jazz Theory Book*. Sher Music. Peta Luma, CA. 183–192

Gradually the limits of the rules are stretched. The in-out technique has become so established that it is regarded sophisticated rather than wrong. However on a basic implicit level, playing outside the harmony is always wrong in the sense that we feel the violation.

Just as with sequences, there is ERP evidence showing that we react to violations of expected structural outcomes such as chord appropriateness. While violations of sequences cause MMN (mismatch negativity), these violations of expected structure cause ERAN (Early Right Anterior Negativity). The ERAN reaction is related to the sense of tonality and the functionality of chords (Koelsch & Mulder 2002)[236]. Contrary to MMN (on sequences), ERAN (structure) reactions do not require prior presentation. It is therefore assumed that ERAN shows that we have an implicit general sense of how it should sound. Although ERAN is more distinct for musicians than for non-musicians, ERAN can be recorded in non-musicians and children (Koelsch & Friederici 2003)[237], indicating that some sense of tonality is developed early in all of us (provided that we are exposed to tonal music). It is a strange and highly illuminative fact that not even this reaction requires attention to the music.

The fact that the MMN for sequence differs from ERAN for rules in many respects (different location, longer reaction time and above all that no prior presentation needed for ERAN) indicates that we are dealing with two different processes. One process concerns a sequence, which is a unique case. The other process concerns a rule.

Listeners have an implicit capacity in detecting and learning structural patterns after a few expositions. Even children can learn this. ERP responses on inappropriate chords have been found in five year olds (Koelsch *et al.* 2003)[238].

---

[236] Koelsch, S. & Mulder, J. (2002) Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology*. Vol. 113. No 6. 862–869

[237] Koelsch, S. & Friederici, A.D. (2003) Toward the neural basis of processing structure in music. Comparative results of different neurophysiological investigation methods. *Annals of the New York Academy of Sciences*. Vol. 999. *The Neurosciences and Music*. 15–28

[238] Koelsch, S., Grossmann, T., Gunter, T.C., Hahne, H., Schröger, E. & Friederici, A.D. (2003) Children processing music: electric brain responses reveal musical com-

ERAN is derived from the ventrolateral prefrontal cortex and the parietal cortex (Koelsch & Mulder 2002). The localization of rule processing in the ventrolateral prefrontal cortex in music has been confirmed by fMRI data (Koelsch *et al.* 2002[239]; Janata *et al.* 2002; Tillmann *et al.* 2003a). The Janata study found indication of a distributed topographic representation of the overall tonality surface in the rostral ventral prefrontal cortex. Interestingly this representation is not absolute; it can reorganize. This allows for a generalization of tonality from one key to another. This is in accordance with the fact that most people are not dependent on a certain key. They do not have absolute pitch.

One year later Janata *et al.* (2003)[240] investigated this representation further. Subjects were presented with melodies played over a variety of scales. Each tonality showed a specific pattern in ventrolateral prefrontal cortex. There is thus a clear indication of this area (BA47) also called the frontal operculum. Prefrontal cortex neurons generally are critical to "the rules of the game" so called "cue-reward" rules or "if-then" rules. This "rule information" is fed back to sensory cortices (top-down processing), acting on the perception of the sensory stimuli (for overview see Chafee & Goldman Rakic (2000)[241].

In 2006 it was reported that the link between the auditory cortex and BA47 is critical for pitch discrimination (Hyde *et al.* 2006). This can be seen in the light of a study (Peretz *et al.* 2005)[242] showing that auditory cortex functions normally among amusical (not able to recognize melodies) subjects. Thus a perfect auditory cortex is not

---

petence and gender differences. *Journal of Cognitive Neuroscience.* Vol. 15. No 5. 683–693

[239] Koelsch, S., Gunter, T.C., von Cramon, D.Y., Zysset, S., Lohmann, G. & Friederici, A. (2002) Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage.* Vol. 17. No 2. 956–966

[240] Janata, P., Birk, J.L., Tillmann, B. & Bharucha, J.J. (2003) Online detection of tonal pop-out in modulating contexts. *Music Perception.* Vol. 20. No 3. 283–305

[241] Chafee, M.V. & Goldman-Rakic, P.S. (2000) Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *Journal of Neurophysiology.* Vol. 83. No 3. 1550–66

[242] Peretz, I., Brattico, E., & Tervaniemi, M. (2005) Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology.* Vol. 58. No 3. 478–82

enough for pitch discrimination. The tones must be related to the representation of tonality (or other possible organizing principles).

Periodicity – rules of repetition and variation

As Peter Kivy (2002) has pointed out, music can be distinguished from narrative arts because music recycles the same material over and over. This would not work in a narrative because it would not push the story forward. Repetition creates periodicity.

Periodicity is a rule. It is very hard to find music that is not periodic. The difference from the sequence is that variation is allowed. It is a rule with many subdivisions. On the implicit level this rule, as the other music rules, work as expected outcomes on cues.

There is an enormous amount of concepts for time in music. I will be very brief and by necessity simplifying. *Beat* refers to where you want to clap or stomp. *Meter* is the amount of beats in one bar, indicated by the time signature in notated music. A bar (or measure) is the distance between the first beat in one bar to the first beat in the following bar. Meter can be played in many different ways depending on style and culture. Rhythm is style specific playing of meter.

The bar is the basic period but there are also layers of longer periodic patterns. These patterns are structured by melody, by chord sequences, by repetition and variation.

The rule is that melodic periodicity coincides with meter. This entails that the tonal centres coincide with this longer periodicity. Most popular music is built by sequences of four bars. There are cycles of four, eight or sixteen bars. This is how long a sequence of chords or a melody arch will typically last before returning to the tonic note and the tonic chord (or a note or chord closely associated with these tonal centres).

However there is an interesting exception from this rule. In the case of *polymetric* music (not polyrhythmic) a melodic figure such as an arpeggio does not even up with the meter. It is called polymetric because you are playing two time signatures at the same time. Simply put, the meter of the played melody does not fit with the tapping of your foot. In the Glen Miller classic 'In the Mood' we can see that a melodic period of three eight notes is presented in a 4/4 meter swing context. This creates the rhythmic characteristic of this song. You can test the difference by clapping the first tone of the three tone melody. Now it is a waltz. And the melody is, although identical, perceived differently. The polymetric technique has been developed by Ralph Towner and others (Towner 1985)[243]. Polymetric rhythms are thus less regular and the sensation is floating and playful. It makes the melody less predictable and thus more interesting.

Just as in architecture, it is the repetition and varied repetition of limited material that constitutes the form. The more limited the material, the more evident is the form. As Johan Sebastian Bach showed, a musical universe can be built from a few elements. Western music exhibits a long list of variations of repeated material (McAdam & Matzkin 2003)[244]:

1. Identical repetition (sequence)
2. Identical repetition but re-orchestration
3. Transposition of pitch
4. Variation of tempo
5. Inversion
6. Retrogradation
7. Variation of the melody
8. Variation of harmonic structure
9. Rhythmic variation
10. Keeping melodic contour, changing intervals

Additionally there are all sorts of combinations of these variations. This list of techniques tells us something important about music: Music is a layer of repetitions and varied repetitions. In popular music the base drum continuously repeats time durations. The base often plays the same figure bar by bar (adapted to the harmony). The harmonies are sequenced in vamps (series of identical chord changes)

---

[243] Towner, R. (1985) *Improvisation and Performance Techniques for Classical and Acoustic Guitar*. 21st Century Music Productions. Wayne, New Jersey

[244] McAdams, S. & Matzkin, D. (2003) The roots of musical variations in perceptual similarity and invariance. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. 79–94

changes) and a funk guitar might present more or less identical four bar "licks" ("ostinatos" for high brow readers). Verses of sixteen bars are repeated and there is an overall structure A-A-B-A-B-C-A-B for example, each letter representing eight or sixteen bars. In classical music the fugue, the rondo, the sonata form etc, are conventionalized patterns of repetitions and variations intertwined with new material.

It is thus the interlocking of the rhythm section with melodic and harmonic repetition and variation that creates periodicity.

To sum up, periodicity

1. creates time structure
2. creates context to forthcoming musical events.
3. makes the music easy to remember (mapped).
4. makes the piece predictable.
5. makes the piece easier to perceive (this follows from the preceding points).

How are musical expectancies created?

As we have seen Gestalt, perceptual learning, sequencing, chunking and music rules are organizing principles that help us anticipate music. Musical anticipations are often "heard". This phenomenon is an aspect of music imagery. Perception and imagery recruit the same areas.

> Those structures responsible for the processing of auditory stimuli are also activated under certain conditions of musical/auditory imagery when no sensory stimuli are present. (Janata 2001)[245]

From this Janata draws an important conclusion:

---

[245] Janata, P. (2001) Neurophysiological mechanisms underlying auditory image formation in music. atonal.ucdavis.edu/manuscripts/Janata_2001c.pdf

> Indirectly, these results suggest that discrete, mentally generated, auditory images might be compared against incoming sensory information in the auditory cortex.

Since imagery can be seen as expectancy, mismatches between imagery and the outcome of a situation may be responsible for ERP reactions on surprising events. Recently this idea has been experimentally confirmed in an ERP study (Yumoto *et al.* 2005)[246]. Yumoto found that imagery generated from score reading causes the ERP reaction called "mismatch negativity" when this imagery is violated by sensory input.

ERAN-reactions on music structures involve auditory cortex and frontal areas. Frontal areas are involved in rule processing. The ERAN-reactions thus indicate that music structure expectancies result from implicit rule processes – in this case rules learnt by perceptual imprinting of music. The incoming stimulus starts generating rule governed consequences.

I suggest that there are two types of musical imagery: Imagery of known music and imagery of new music. The first type engages the sequence circuit and the second engages the rule circuit.

ERPs can be derived from violations of long term memories and of violations of short term memories. From this follows that some imagery engages short term memory while other imagery engages long term memory. Thus we would have four possible categories of music imagery:

1. Short term memory – sequence processing (e.g. the imagery of a melody just heard)
2. Short term memory – rule processing. (I have no example of this since I doubt that implicit music rules can be learnt in one hearing)
3. Long term memory – sequence processing (the imagery of a known song)

---

[246] Yumoto, M., Matsuda, M., Iteoh, K., Uno, A., Karino, S. Saitoh, O., Kaneko, Y., Yatomi, Y. & Kaga, K. (2005) Auditory imagery mismatch negativity elicited in musicians. *Neuroreport.* Vol. 16. No 11. 1175–1178

4. Long term memory – rule processing (imagery of new music)

The last case can be a felt continuance of a piece never heard before (completion), but it can also be the experience of hearing a whole new piece of music being played in your head.

What is causing imagery on a neurological level? The cerebellum feeds forward to predict sequences and feeds back for error correction. One of the functions of cerebellum is thus to detect, prevent, and correct mismatches between intended outcome, and the perceived outcome, of interaction with the environment (Ito 1993)[247]. This role of cerebellum is supported by the observation that cerebellar disorder causes impaired contact with reality.

According to Jeannerod & Frak (1999) the cerebellum initiates action sequences, based on past learning. These commands reach the motor cortex, where they are partly inhibited by prefrontal signals. The implication is that the prefrontal area selects an alternative fitting rule. This would, according to the authors, generate action imagery. Cerebellar sensori-motor driven imagery thus is an important part of the perception process. How does this work?

When the organism detects a stimulus on a basic brainstem thalamic level, this initiates cerebellar activity in order to direct attention to the stimulus (Schmahmann *et al.* 2001)[248]. This looking-for activity triggers sensorimotor imagery with possible action affordances before the object is attended to. Thus the brain already has a vague expectancy as it directs its attention. From this we can draw two conclusions:

1. The involvement of motor areas confirms that perception to a large extent is to act. To listen to a sound is to latently produce it.

2. Imagery has, beside the function to fill out sensory income with "suggested" content, the function to direct attention to

---

[247] Ito, M. (1993) Movement and thought: Identical control mechanisms by the cerebellum. *Trends in Neurosciences*. Vol. 16. No 11. 448–450

[248] Schmahmann, J.D., Anderson, C. M., Newton, N. & Ellis, R. (2001) The function of the Cerebellum. *Consciousness & Emotion*. Vol. 2. No 2. 273–309

unattended events. It can produce "the taste of danger" before the cause of the danger is identified.

Discussion

The overall impression from the literature is that our mental equipment for perceiving the world (Gestalts, perceptual learning, sequencing and rule processing) is reflected in music structure. In other words – music structure mirrors the human mind. It is obvious that musicians over the years have developed techniques that make musical structures easy to perceive, remember and predict. This is not so much the result of the composers' intentions as the pragmatic finding that some techniques seem to work. They make music user-friendly. Popular.

The ERP studies show that music structures are mapped as schemas in our brains. These schemas are activated by music. As we listen we follow this schema in real time. The sensory input needs to be related to this mapped structure to be perceived. MMN data suggest that auditory sensory input is "pre-perceptual, fragmentary, lacks the time dimension, and cannot enter conscious perception". It is not until it "is mapped onto the neurophysiological basis of sensory memory in the auditory cortex" that we become aware of the perception (Winkler and Näätänen 2006)[249].

This confirms the idea that the sensory input does not produce meaning until it triggers imagery.

I have presented some mechanisms underlying the perception of music. They provide us with a *local view* of the topography. The concept "local view" is the cue for the second part of this chapter.

Principles of navigation

Navigation is the art and science of determining one's position so as to safely travel to a desired destination. This art uses allocentric

[249] Winkler, I. & Näätänen, R. (2006) Theoretical framework of pre-attentive auditory processing. http://www.cbru.helsinki.fi/research/theory/description.html

landmarks such as the stars, the North Pole, the sun, lighthouses, fog-horns or stationary satellites. The terrain is mapped by triangulation of allocentric landmarks.

But navigation is also a sense universally shared among animals. This is, at least among mammals, a similar mechanism across species. The study of navigation in animals is a niche of experimental research. The subjects are usually mice running in mazes. From these studies the researchers draw conclusions concerning human navigation.

David Redish discerns five models of navigation strategies in the literature: random navigation, taxon navigation, praxic navigation, route navigation and local navigation (Redish 1999)[250]. Of these four, two are relevant in our comparison: Praxic navigation and local navigation.

Praxic navigation is thought as a motor program: if the animal starts from the same place in the same direction, procedural memory tells the animal where it is. Once the sequence of movements begins it completes to the end without external input. The animal learns a sequence of turns and directional movements that will take it to the goal. Once the maze is learned, praxic navigation can be used in the dark. When the maze is darkened the rat has to change the frame of reference from allocentric landmarks to egocentric landmarks – from external patterns of structure to internal patterns of movements. Instead of relying on visible cues, it now relies on motor sequences of running and turning. The movement is automated. Praxic navigation is procedural.

Local navigation demands cognitive maps – a coordinate representation of the location. The anatomical locus of this map is the hippocampus. The map consists of allocentric landmarks. Local navigation is according to Redish based on

1. *the local view*: a representation of the animal's relationship to the landmarks in its environment

---

[250] Redish, D. (1999) *Beyond the Cognitive Map – from place cells to episodic memory*. The MIT Press, Cambridge, Massachusetts

2. *path integration* coordinates: a metric sequenced representation of position (map) that accommodates vector arithmetic
3. *head direction:* a representation of orientation in space
4. *place code:* a distributed representation of position that ties local views to path integration coordinates
5. *goal memory:* a representation that associates motivational and spatial inputs and accommodates trajectory planning

I will now suggest parallels between local navigation and the perception of music structure. The suggested applications are put in brackets:

1. The local view is the present moment perception of the piece (music momentary heard).
2. Path integration. As may be recalled from Chapter 1, taking a perspective is to integrate a perspective with the egocentric perspective. Local navigation theory fits this description in that the allocentric perspective is integrated with the egocentric perspective. Redding describes path integration as the integration of vector arithmetic with map position. Since rats do not do arithmetic I take this to be a metaphor for the feedback of own movement. This movement is integrated in the allocentric map (finger movements + music structure).
3. Head direction is not relevant for place detection in music structure.
4. Place code is described as a distributed representation (tonality!) of position tying local view (the music momentary perceived) with integration co-ordinates (tonality + finger movements).
5. Goal memory is not relevant. The task in music is not to get to the goal. In music as in life it is the road that makes

> the effort worthwhile. We may however have sub goals such as the return to the tonic.

Locale navigation could be described as an integration of local view, the own movement feed back and map representation. Two types of memory are relevant for navigation: the procedural memory for praxic navigation, which is the body's memory how to move, and the allocentric geographic map of the landscape (Berthoz 2000).

Navigation in music

Crossing a field generates motor/tactile, visual, and acoustic feedback. The path is a sequence of turns and crests, a sequence of related body movements and a sequence of feedback sensations. As the path is learned – representations of these patterns become interconnected.

The instrumentalist has to know where she is and she must have the ability to go from one point to another in the music. She does so by movements on the instrument. This is similar to driving a vehicle – she moves vis-à-vis the instrument and this takes her to the destination.

On musical instruments every key represents a tone. As the fingers of the pianist travel over the keyboard she is travelling in the music. Every sequence of sounds is a sequence of places.

To play an instrument is to move the keys to achieve acoustic feedback.

As we learn to play, we use allocentric landmarks. We laboriously translate heard notes or notes from the sheet to the keys on the instrument. Gradually the body learns the movement pattern. The visual or heard structure is connected to movement structure – a bodily representation. When automated, the sequences in the score, the sequences of the keys, the sequences of sound, and the sequences of body movements (motor/tactile feedback) are locked to each other. Now we move the body through the musical structure. We have integrated the allocentric perspective with the egocentric perspective. From now on we can use praxic navigation. Just like the rat in the maze, the player can perform in the dark. When the piece is learnt, musicians do not have to look at the sheet or the instrument to play.

Can musical topography be compared to the landscape? McAdams & Matzkin (2003) discern three main types of musical invariance:

1. Similar texture
2. Figural similarity
3. Similar structural abstraction

If we allow ourselves to play with the resemblance of perception of music to the perception of the landscape, we may see that similar texture (orchestration, ornamentation, interpretation etc.) is like the illumination of the landscape, the colour, the season etc.; figural similarity (melody) is the paths and its contours; and structural abstraction (harmony and form) is the underlying structure in spite of changes in the illumination, the season and the path.

Now imagine that you are walking in the landscape and we have added rhythm.

The path

As listeners we can choose to follow one instrument. Usually this is the melody. More advanced listeners can chose to follow the base line or even intertwining lines in a polyphonic musical weave. As we do these choices we chose different perspectives within the allocentric perspective, since we foreground one line and background the others. A familiar piece of music is mapped as a fixed sequence, whereas styles and genres are maps with forks.

A jazz musician has a set of pre-programmed sequences that she plays over a given chord succession. Here and there she comes to the end of a sequence and has to choose from a set of directions given by the musical context (cue). She knows the place. Some directions are possible and more or less probable and some are clear violations of the rule.

Listening to this player is like following her through a town where every corner offers three solutions but not a fourth or fifth solution.

The experience of creativeness in music is to come to a fork and find the fourth solution – a passage nobody has seen before leading to a familiar place. For the listener it is an adventure to follow this – Yeah, yeah, OK, hm? Aha! Yeah, What? OK – and so on. Even if the road is not new to the musician, it might be to the audience.

Few musicians are creative in the sense that they find new ways during improvisation. Rather they have tested different solutions during practice. These solutions present themselves to the musician during the improvisation. There may not be an entirely new melodic solution on some jazz vamps any more. The originality is on another level. It is like chess. There is not a variant in chess that has not been tested at least 20 moves ahead. Originality is about playing the less usual and combining in new ways.

As the listener moves along in the music, every mismatch is contributing new information to the context and changes the hypothesis of what is actually heard. The sense for probable outcomes is a sense for music, or musicality, if you want.

This is in line with Neisser's idea (1976) of the perception as a continuous cyclical process, which describes perception as a series of rapid hypothesis testing. Mismatches continuously change the expectancies. They violate the sense of style. When the mismatch is worn out, we have a new style. This is then a neurological comment on Leonard Meyer's discussion on changes of style as a result of how the unexpected becomes norm (Meyer 1956). However, luckily it seems that the mismatch reaction is difficult to erase. If we get rid of them totally, music would loose its meaning. We have to be able to discriminate the feeling of dissonance from the feeling of consonance, the feeling of tones outside the harmonic context from the feeling of tones inside this context, the feeling of syncopation from on beat events, and so on.

Landmarks

We have stated that the perception of structure demands landmarks and that those landmarks are called allocentric landmarks for the terrain. What is the counterpart for music? I argue that the tonal centres – *the tonic note* (e.g. the tone C in the key of C) and the tonic chord (e.g. the C major or minor chord in the key of C) serve as landmarks. They differ, however, from the usual understanding, since they are not located in spatial structure but in sound structure. Can we rightfully talk about allocentric landmarks in the case of music. I will argue that we can.

These landmarks are allocentric in the sense that they are outside the body of the listener. The tonic note provides a cognitive reference point, a "melodic anchor" (Bregman 1990). We feel when we are there and when we are not. All the other notes in a melody have some kind of felt potential difference to the tonic note. The tonic note, or closely related notes (the third or the fifth), is often the first and the last points in a melody. It is a point of departure and a point of arrival. We can feel this arrival as well as the expectancy of arrival.

Krumhansl (1995)[251] found a psychological hierarchical organization of the seven pitches of the diatonic scale around a reference tone. By measuring the psychological distance between tones occupying different positions in this hierarchy he showed that the stable tones serve as reference points in relation to which all the other tones are perceived. The tonic note is in other words a landmark in our perception of music.

The same could be said about the tonic chord. It is the most expected chord. Sequences of chords tend to lead to the tonic. We know this but it has also been investigated statistically (Huron 2006).

Tillmann, Bharucha & Bigand (2003b)[252] built artificial neural network (ANN) models of *SOMs* (Self Organizing Maps). The SOMs were programmed to learn from recordings of music pieces. From this learning SOMs became capable of localizing tonal centres.

Humans have been given the same task as the machines (to find tonal centres). A cross-cultural study in the 80s revealed interesting differences (Kessler, Hansen & Shepard 1984)[253]. If the subject is not

---

[251] Krumhansl, C.L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*. Vol. 17. No 1. 53–80

[252] Tillmann, B., Bharucha, J.J. & Bigand, E. (2003b) Learning and perceiving musical structures: Further insights from artificial networks. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. Chapter 8.

[253] Kessler, E.J., Hansen, C., & Shepard, R.N. (1984) Tonal schemata in the perception of music in Bali and the West. *Music Perception*, 2. 131–165

enculturated to the organizing principle of the piece she is exposed to, she will tend to say that the most frequent tone in the piece is the tonal centre. If, on the other hand, she is familiar with the organizing principle, her judgement will be directed by hierarchical choices governed by the principle. This shows that the organizing principle is recognized if once learned and that this principle guides expectancies and thus the perception of music. The importance of this study is that it proves that the sense of tonality is not only based on expectancies grounded on how frequent a certain tone is, but that there really is a rule-mechanism governing this sense.

The sense of tonality does not require explicit knowledge. Even if the listener is not educated in music and does not know what a tonic is, she would, if asked to continue a given melody, continue according to the rules of tonality. The only condition for learning the rules of tonality is that we are exposed to tonal music.

Who remembers atonal music? In the 1920s continuity was considered a bourgeoisie fetish by radicals. Freud exposed cultural hierarchies as emblems of oppression. The avant-gardism movement broke up the hierarchies; art turned non-figurative; poetry ridded itself of rhymes and meter and in music Arnold Schönberg constructed the homogeneous twelve-tone scale. Every tone had the same value. The hierarchy was broken. Theodor Adorno saw this as revolutionary music. No tone submitted to another.

If the scale is homogeneous there cannot be any tonal centre. This is why this music is called atonal. And this is why most of us get lost in it. It is like a landscape without landmarks. If the landscape is totally homogeneous, like the open sea or chaotic like a snowstorm, we cannot find our way home. We are reduced to use random navigation. As Claude Lévi-Strauss (the father of structuralism) put it:

> [Serialism] is like a sailless ship, driven out to sea by its captain, who is privately convinced that by subjecting life aboard to the rules of an elaborate protocol, he will prevent the crew from thinking nostalgically either of their home port or of their ultimate destination (Claude Lévi-Strauss 1975: 113)[254]

[254] Lévi-Strauss, C. (1975) *The Raw and the Cooked. Introduction to a science of mythology.* Harper, New York

There are rules in atonal music, but the average concert audience has no access to them. They are not internalized in the body of the listener. It is a theoretical construction. The sounding result of these calculations may come as a surprise even to the composer (as I happen to know). It is not user-friendly. It does not create expectations as other music typically does. This protocol obstructs us from knowing where we are and where we are heading. It is made that way. The human brain is unequipped to perceive structure without landmarks. Since expectancy is, according to some musicologists, Huron to name one, a prerequisite for positive musical emotion, we might ask if it is possible to enjoy atonal music. This has not to my knowledge been tested.

Fred Lehrdal (1988)[255] listed the cognitive constraints on compositional systems. Constraint No 2 reads: "The musical surface must be available for hierarchical structuring by the listening grammar".

The brain, always trying to make something out of the input, reacts. The sense of chaos is alarming. In films atonal music typically signifies loss of control, insanity and horror.

The musical hegemony has gone through some strain to teach and distribute this music (not enough according to its composers). With this exposition it ought to be liked by means of perceptual learning. After all these years, it still is not requested. Ten seconds is enough to perceive such an object. Hours of listening do not add more if we cannot follow the development.

Nowadays atonal music is merely a tool in the toolbox – comparable to other tools such as counterpoint, according to the Swedish composer Anders Hultqvist (Hultqvist 2007)[256]. He describes contemporary music as a continuum from atonal music to consonant triads. The technique is usually mixed with other techniques. It creates effects by means of contrast.

We have already stated that time must be one of the dimensions in the integrated allocentric/egocentric perspective.

---

[255] Lehrdal, F. (1988) Cognitive constrains on compositional systems. *Generative Processes in Music.* (Ed. John Sloboda). Oxford University Press. 231–259

[256] Hultqvist, A. (2007) Personal information 20/4

I suggest that the "One" – the first beat in each bar – serves as a temporal landmark. It is easy to perceive and the rest of the events in the bar are related to it. Particularly strong landmarks are those "Ones" that mark the beginning of a period. These are often indicated by a change of tonality, or a change of instrumentation. Furthermore they are often announced by fills and cadences.

Music thus is hierarchically structured in respect to tonality and periodicity. The rhythm, the chords and the melody make up a musical topography. The musical topography is created from landmarks *in* the music structure – allocentric landmarks.

Now let us compare navigation and music listening on a neurological basis.

Neurological level

Separate streams

The ventral pre-motor area in the brain is connected to commands for grasping which demands perception of the form of the object and the dorsal pre-motor area is connected to reaching which demands a sense of direction and movement (Wolfensteller *et al.* 2004)[257].

This functional segregation is reflected in the perception process. Visual information is thus processed along one ventral stream and one dorsal stream.

The ventral stream processes object identification (needed for grasping) and localisation in allocentric space. The dorsal stream processes localisation in egocentric space. It triggers fast procedural online reactions (motor response).

Only ventral visual information can be reported. This means that although dorsal information is processed, it is pre-attentive. Consequently patients with impaired ventral stream report that they cannot

[257] Wolfensteller, U., Schubotz, R.I., & von Cramon, D.Y. (2004) "What" becoming "Where":
Functional magnetic resonance imaging evidence for pragmatic relevance driving pre-motor cortex. *The Journal of Neuroscience*, Nov. 17. Vol. 24. No 46. 10431–10439

see objects. They cannot describe form or colour. If the dorsal stream is intact, they may still be able to act on visual information. This is called “cortical blindness” or “blindsight”. Patients diagnosed with “blindsight” can point out directions to objects even if they report that they cannot see them. What’s more, they can reliably perceive facial expressions and possibly react emotionally on body language (de Gelder & Hadjikhani 2006)[258].

Burgess (2002)[259] suggests that the two streams meet in hippocampus and parahippocampus. This is thus a possible location for the integration of the egocentric and the allocentric perspectives. It is according Berthoz (2000) only in the hippocampus that the body is placed in *allocentric space* (space organized around allocentric landmarks).

Pre-motor cortex for sound processing has a corresponding ventral/dorsal organization (Zatorre *et al.* 2007). Jääskeleinen *et al.* (2004)[260] report evidence from ERP-experiments suggesting segregated “What” (ventral) and “Where” (dorsal) auditory streams. This was confirmed in an fMRI study testing for functional connectivity of the processing of word sequences (Ruchsbaum *et al.* 2005)[261]. The authors distinguished a ventral stream from the superior temporal gyrus to ventrolateral prefrontal gyrus and a dorsal stream from the planum temporale (in the auditory cortex) to the dorsolateral prefrontal gyrus. The ventral pre-motor area thus reflects sounds as direct motor imitations, such as vocalization and grips for chords, whereas the dorsal pre-motor area is indirect and rule based.

The ventral stream processes present now perception. It is accessible for consciousness and can be analysed for meaning (such as source identification).

---

[258] De Gelder, B. & Hadjikhani, N. (2006) Non conscious recognition of body language. *Neuroreport.* Vol 17. No 6. 583–586

[259] Burgess, N. (2002) The hippocampus, space, and viewpoints in episodic memory. *The Quarterly Journal of Experimental Psycholology.* Vol. 55. No 4. 1057–1080

[260] Jääskeläinen, I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levänen, S. Lin, F., May, P., Melcher, J. Stufflebeam, S., Tiitinen, H. & Belliveau, J.W. (2004) Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the New York Academy of Sciences.* Vol. 101. No 17. 6809–6814

[261] Ruchsbaum, B., Olsen, R., Koch, P. & Berman, K. (2005) Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron.* Vol. 41. No 4. 687–697

The dorsal stream is faster and multimodal (vision + hearing). It processes heard speech, as well as read text. It connects the planum temporale with prefrontal areas. The planum temporale has been identified as an interface between perception and the production of sounds (Hickok *et al.* 2003). It is associated with articulation and action planning. The planum temporale is considered a gateway to frontal motor systems. It is associated with all kind of sequential patterns (speech, music, rhythm). The planum temporale reacts to unexpected durations in sequences (Mustovic *et al.* 2003).

As may be recalled Janata located a prefrontal area, BA47, sensitive to tonality. This is one of the prefrontal areas where the pathway from the planum temporale connects in the case of music. BA47 is a rule processor of music.

Melodies and harmonies activate the planum temporale on the right temporal lobe, whereas the speech centre is to be found on the corresponding left side (Peretz 2003)[262]. Both these areas are thus involved in syntactic processing. Bermudez & Zatorre (2005)[263] found that the right planum temporale has higher concentrations of grey matter for musicians than for non-musicians. As stated before, the processing of tonality and pitch is dependent on the connection between the planum temporale and BA47 in prefrontal cortex.

This general picture of the dorsal stream processing of sound is supported and completed by Koelsch (*et al.* 2006). The authors detect "a motor-related circuitry for music comprising the ventral striatum, the insular cortex, the Rolandic, and possibly the frontal operculum". The frontal operculum BA44 detected active in this study is generally involved in linguistic and music rule and sequence processing (Koelsch 2007)[264]. This area is adjacent to BA47 (tonality rules). This indicates a general frontal lobe area divided in sub areas designed for different syntactic tasks.

---

[262] Peretz, I. (2003) Brain specialization for music: New evidence from congenital amusia. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. 192–203

[263] Bermudez, P. & Zatorre R.J. (2005) Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences.* Vol.1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo). 395–399

[264] Koelsch, S. (2007) Personal information 7/7

To sum up, the evidence indicates that music is perceived along segregated streams:

1. A ventral "What-stream" for identification of sound, accessible to consciousness. This stream produces pre-motor imitation of sounds.
2. A dorsal "Where-stream" that is pre-attentive and rule based. This stream is insensitive to modality and organizes the information along sequential, rhythmical and tonal parameters.

These streams correspond to the egocentric (identification) and the allocentric (rule based structure) perspectives.

Integration of streams

As already has been pointed out, it has been proposed that the ventral and the dorsal visual streams converge in the hippocampus and the parahippocampus (Burgess 2002). We have also seen that the ventral stream and the dorsal stream may be associated with the egocentric and the allocentric perspective. "Hippocampus allows the animal to self-localize on re-entering a familiar environment. This allows the animal to reset its path integrator representation from the local view" (Redish 1999). I suggest that what Redish calls "the path integrator" is the integration of the map (allocentric perspective) with the egocentric perspective. When integrated, the animal lives the map from its own perspective. As this happens the position of the body is related to the allocentric landmarks and the sequence of the route appears. A sequence of memories is replayed. It involves the path and all the cues on the path. This must be so because the path is defined through the cues. Since these cues are lived emotionally, the total experience is replayed as lived – path and context.

Hippocampus is involved in the transfer from *short-term memory* to long term storage (Gazzaniga 1985). It is well known that hippocampal injury causes amnesia. It has also been noted among amnesia patients that new information does not seem to have emotional value.

The sequences of place cell activation in the rat hippocampus during REM sleep mirror the sequence when the rat is navigating in a maze. The neurons learn the order to fire. This order can never be reversed. Even the neurons "beside the path" fire but weaker. A vector field is thus created in hippocampus, pointing to the goal (Redish 1999). The episodic memory is such a replay. What is recalled is the situation as lived. If, for example the animal once has been frightened when it passed a rock, this rock will forever be imprinted emotionally as a dangerous place.

The irreversible vector field for episodic memories of paths may be a neurological explanation why we cannot remember music backwards. If time is abstracted from the memory we may semantically be able to name a music piece, but we cannot recall it without the time dimension. This replay may explain music imagery (inner music perception not triggered by sensory input). This puts a cognitive constraint on music. It is possible that most listeners cannot perceive retrogrades (reversed order of tones) of themes. It is not surprising that this technique cannot be found in popular music. It does not sell records.

Hippocampus has been found active in contextual shifts. This means that hippocampus is involved in allocentrically cued shifts. Music structure provides contextual associations within the piece by means of repetition, and between pieces by means of recognition of style and genre. I suggest that contextual shifts in music, or violations of contextual expectances in music, imply hippocampal activation. To put a rat on a familiar track may in some basic neurological respect be comparable to putting on a familiar record for a listener. The contextual representation, be it music context or terrain context, is reset at the presentation: "Where am I? I recognize this; I might be here. If so, I expect this and this, and this. Match! I am here! Now I expect this." These questions are a pre-attentive part of the perceptive cycle. We will come back to the emotional implications for music.

I suggest that the auditory streams just as the visual streams converge in hippocampus and parahippocampus. The reason is an fMRI study, which tested where the brain processes violations of harmony and violation of expected timbre. Although both tests showed various activations practically all over the brain, only one place showed interaction between timbre (ventral) and structure (dorsal) expectations: the parahippocampus (Tillmann *et al.* 2006)[265].

Even if the information is integrated in the hippocampus and the hippocampus is associated with long-term memory this does not mean that all memories are stored in the hippocampus. Rather the information is fed back to the perception and association areas. Episodic memorizing is to reactivate the perception process. This is what makes episodic memories so vivid. The process can be activated by any contextual cue. Since we live the event anew, this is by definition imagery. I believe that imagery and episodic memory is connected.

We can actually see examples of the integration of the streams. In televised music quiz games, where the contender is asked to remember the title of a popular song (contextual information), you can clearly see how she starts moving into the song, snapping fingers, waving the body, in order to recall the lyrics. Technically speaking this could be analysed as an attempt to set the path integrator – to coordinate the body with the music track. The lyrics are actualized through the body. When the contender finds a word, the path is integrated and the sequence of the words in the melody will come to her. Then she sings on until she finds the word that cues to the title. This strategy is spontaneous. The example shows that we may indeed use praxic (bodily) navigation strategy to arrive to contextual knowledge.

Another example of music-context connection can be observed in people suffering from stroke who have lost their speech. They can often sing song texts.

We can conclude from the evidence and the knowledge at the present state that music is processed along a dorsal stream connecting auditory cortex and specifically planum temporale with motor areas and with BA47 (the rules of tonality and sequences) and a ventral stream for identification of sounds and sound sources. These two pathways possibly converge in the hippocampus. The hippo-

---

[265] Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D. & von Cramon, D.Y. (2006) Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage.* Vol. 31. No 4. 1771–1782

campal area will associate the music with context such as styles, places, lyrics and emotions. However these emotions will be elicited by contextual associations.

Neurological correspondence between place and music

Experiments have revealed that rats have "place cells" in the hippocampus firing each time the rat passes a certain place (Berthoz 2000). The rat thus has an allocentric topographical representation of its natural environment in the hippocampus. There is evidence for a neural code of human spatial navigation based on cells that respond at specific spatial locations and cells that respond to views of landmarks (Ekstrom *et al.* 2003)[266]. This study identifies place cells (primarily found in the hippocampus), as well as "view" cells (responsive to landmarks; found mainly in the parahippocampal region) and "goal" cells (responsive to goals, found throughout the frontal and temporal lobes). Some cells respond to combinations of place, view and goal — for example, cells that responded to viewing an object only when that object was a goal.

You may get the impression that allocentric landmarks in the terrain are mapped in hippocampus whereas the allocentric landmarks in music are mapped in BA47 in prefrontal cortex and from this draw the conclusion that the perception of music and the perception of a landscape are basically different. However tonality must be learnt, since tonality is not a universal phenomenon. There can thus not be a brain area exclusively designed to process tonality but rather an area equipped to learn the rules which can be applied for tonality. Can this area be used for cueing for place? Area BA47 shows spatial selectivity (Rizzuto *et al.* 2005)[267]. It is critical for visuo-spatial movement planning. And furthermore, the hippocampus

---

[266] Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T., Isham, E., Newman, E. & Fried, I. (2003) Cellular networks underlying human spatial navigation. *Nature.* Vol. 425. No 6954. 184–187

[267] Rizzuto, D., Mamelak, A., Sutherling, W., Fineman, I. & Andersen, R.A. (2005) Spatial selectivity in human ventrolateral prefrontal cortex. *Nature Neuroscience.* Vol. 8. No 4. 415–417

together with the caudate nucleus and amygdala is activated by chords violating harmonic relationships (Tillmann *et al.* 2003). This shows, according to the authors, that there is a tonal representation in the hippocampus possibly mapped by place cells. Thus there is one area known to map music that also processes location (BA47), and another area known to map location that also processes music (the hippocampus). This indicates a close relation between the perception of musical and spatial structure.

Implications

The phenomenology of technique

The integration of the allocentric perspective with the egocentric perspective entails that structure is translated to movement. The music sequence is translated into a movement sequence, just as the sequence of the path is translated into movement. Since the planum temporale connects tonality representation in the prefrontal cortex with motor activity, we can see the planum temporale as the interface between structure and movement. The movements of the player give rise to music structure and the structure of the music causes movements. This interface could be spelled *technique.*

Technique is usually defined as the performance of a task with a minimal loss of energy. Since economizing the energy is very important to most creatures, we can expect that animals have a sense for technique. The bird on its way over open sea economizes to the limit. The cheetah balances the cost of energy with the gain of the hunt. For the sparrow surviving the winter, every move must contribute to nourishment. The movement patterns of all animals are part of this cost-benefit equation. This indicates that technique is a sense. When we do something often, we tend to find this perfect balance between aim and cost. There is thus an ultimate movement pattern for each task. This could be demonstrated by specifying the animals as slalom skiers competing in a race. The structure is decided by the position of the gates. Top skiers tend to move identically. Even the line between gates is identical. There is for each skier an optimal way to perform the race. This movement is a "compromise" between the path, the skies and the body. If the path is fixed, the optimal choreography is fixed.

The movements become automated. The technique, learned by imitation, videos and practicing, is procedural knowledge. When the skier starts, she has to rely on this knowledge. This knowledge is automatically adjusted to the terrain. All the accumulated and generalized implicit knowledge from imitations and thousands of previous runs is used in the special case.

If the skier has mapped the track, the movements will follow. Anja Pärson, super-G champion, in Alpine World Championship, Bormio 31/1 2005 has expressed this:

> The excitement makes me tremble like an aspen leaf when I approach the start, and I get so tense by the pumping adrenalin that I sometimes feel as if I had lactic acid even before I have skied one meter. But when it's time to race I get tunnel vision. I see where the line goes from gate to gate and almost have no time to think during the race. (Anja Pärson 2005)[268]

This is a description of total focus. The autopilot turns on. The skier is locked to the target. This state of consciousness is necessary to handle situations at 115 km/h. If Anja started thinking about her movements she would loose accuracy and probably end up hanging in the safety net.

Music performance can be compared to skiing. Two players performing the same melody on the same type of instrument in the same tempo will have to move identically. The first violinist section in a symphony orchestra moves like a shoal of fish. In this case the technique is a compromise between the music, the instrument and the body. But since the body schema includes tools, the instrument (as the skies) is perceived as a body part. The technique for the violinist is a compromise between the sound structure and the body schema. Technique in music, thus, is an "agreement" between the body, and the structure to be followed during the course of time. Speed and accuracy excludes tension and inaccurate movements.

---

[268] Pärson, A. (2005) Interview in *Göteborgs-Posten.* http://www.gp/jsp/crosslink

Visible cues from the score or audible cues from the music are gradually mapped as finger movements. Any improvising musician has a set of "licks" (melody fragments) suited for musical structures, such as scales or arpeggios (chord tones), automated as sequences of finger movements. The player has a set of melodies in her fingers. This procedural knowledge is automatically adjusted to the particular tune (as the technique of the skier is adjusted to every slope).

Charlie Parker used a surprisingly limited set of basic patterns in his improvisations. These patterns could be used in most of the music he performed. Parker adjusted these patterns not only to harmony but to the melody of the original song as well (Martin 1996)[269]. The patterns may be consciously learnt, but they are not consciously used. Rather they are automated sequences of movement connections cued to music feed back.

The task in music playing is to produce perfect pitch, timing and sound. The professional player tries to achieve the ideal technique to do so. This imagined goal is never fully achieved but she closes in on it during years of training. For the professional player the sense of technique is not enough. The work with the technique is about becoming aware of every aspect of movement-sound connection and then to rehearse the movement pattern until it is automated. It happens that musicians discover that they are making some fundamental technical error. Flutists can develop a kind of lip-tongue neurosis, akin to golf swing neurosis, which forces them to disassemble the pattern and assemble it anew. When the technique works the player feels control and this interaction with the instrument is confirming her *musical self* as a co-perception of the sound.

How does the technique of the artist affect the listener? If the listener is a musician she may become something of a "back seat driver", "helping" the musician, since the sound is automatically connected to her body. In a concert this imitation is helped by visual cues. As listeners, we transmit this accuracy to our own bodies. This has been treated in the dyadic chapter.

Luckily the audience is unaware of the pain of practice. Good technique is recognized as seemingly effortless performance. It

[269] Martin, H. (1996) *Charlie Parker and Thematic Improvisation.* Scarecrow Press, Metuchen N.Y.

seems effortless because there is no ineffective strain or tension. It is perceived as elegance, perfection – or even grace. The experience of technical performance is an aesthetic impression in its own right and as such part of our enjoyment of music. Seeing an artist perform with perfect technique is inspiring. Just to have experienced an Andalucian flamenco artist once, may influence a guitarist for life.

To sum up:

1. Technique is a compromise between the structure and the body.
2. The structure of the task determines the technique.
3. The optimal technique is maximal task accuracy combined with minimal effort.
4. Technique is the conciliation of the allocentric and the egocentric perspective.
5. In navigation this conciliation seems to take place in the hippocampus.
6. In music listening we do not know the locus for this integration. The dorsal stream connects the rules of tonality with perceived (unattended) sounds and movements and the ventral stream identifies sounds and sound sources. I speculate that these streams converge in the hippocampal area.

Is musical perceived visually?

As stated in the depiction of the allocentric perspective, visual perception of place is not a prerequisite for this perspective. We can hear place in music. Do we visualize those places?

The integration of the musical streams in the hippocampus allows for contextual associations to music structures. One such possibility is the connection between musical place and geographic place, that is, between heard place and visualized place. Do we perceive place as the music structure unfolds? To some extent this question is a question about cross-modal connections between hearing and sight, since sight always entails spatial place. Such connections exist. The visual and the auditory pathways partly overlap (Hall 2003)[270]. We have already seen that the dorsal stream is cross-modal, that BA47 codes for tonality and place and that the hippocampus is activated by music processing as well as place processing.

Do we perceive visions of places as we listen to music? Do we for example perceive temporal and pitch relations as length and height. I think the answer to this question is yes – sometimes. Here are some examples:

1. We say that melody tones or chords are close to or distant from each other. Do we really perceive this? Or is it a just a metaphor. Surely we feel likeness and unlikeness. However according to Lakoff & Johnson (1980), metaphors are grounded in perception. The choice of metaphor indicates that closeness and distance can in fact be perceived in music. This is embedded in our perception of the world. When we perceive the same structure, we may feel that we have come back to the same place. And, furthermore closeness in time is often perceived as closeness in space, since it takes time to move from one place to another. Thus it can be argued that there is a cognitive relation between the perception of music and the perception of spatiality.

2. If the listener knows how to read music, she can link the heard structure with the written note system. Remembering ordinary text engages allocentric judgement, since we can remember where on the page the word is and the spatial distribution of the letters in the word. There is a functional equivalence to spatial representation and reading. Reading a music score also requires allocentric judgement, and even more so, since the music score is a graph of the sounding music. This means that if the listener is a

---

[270] Hall, D.A. (2003) Auditory Pathways: Are 'What' and 'Where' Appropriate? *Current Biology*. Vol. 13. No 10. 406–408

musician the melody heard might produce an imagination of the melody in writing which would put her on an allocentric map.

3. Most of us, irrespective of if we can read music or not, perceive notes high in pitch as spatially high and vice versa. And to a lesser extent we perceive music as moving from the left to the right simply because we are used to reading ordinary text and graphs in that direction. The time line in Western culture is always a left to right representation. This means that the music is perceived in a two-dimensional space – a horizontal time dimension and a vertical pitch dimension.

4. Instrumentalists associate heard music with places on their instruments.

These first four points are about place perception directly related to the music. However music often triggers or connects to other visual experiences:

1. Music propels visual imagery. In the Gabrielsson & Lindström Wik (2003) material there are numerous reports of music triggering visions.

2. Shamans all over the world use music as a means to travel spiritually.

3. Laboratory-based studies show unequivocal evidence for cross-modal links in spatial attention (Spence & Read 2003)[271].

4. The effect of music listening on the performance of driving in computer games has been tested (North & Hargreaves

[271] Spence, C. & Read, L. (2003) Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *American Psychological Society*. Vol. 14. No 3. 251–256

1999)[272]. The authors suggest a direct link between music and the listening context (driving).

5. Music is used in moving pictures (films, computer games, and commercials) to support movement.

Let us stay for a moment at this last point! Why do music sequences go so well with visual sequences? We have on one hand a heard structure, and on the other hand a seen structure. Both are unfolding over time. I would say that the connection between the two will preferably be made if there is some kind of isomorphism at hand, that is, a structural likeness.

Marshall and Cohan (1988)[273] developed a model for the perception of film based on their research on film music. This model is called the congruence-associationist model. The authors argue that watching a movie, our attention will be drawn to the cross-modal congruence between the music and the moving images. This means that if there is some form of structural likeness between the music and a part of the visual field, this part is likely to be attended to. Music, they claim, directs the eye to visual parts that are structurally congruent with the music.

The connection between the flow of images and music indicates that there is indeed some common underlying mechanism. In other words there is an affinity between the perception of structure in music and the perception of the structure of visual environment.

I suggest a phenomenological explanation for the sound-image connection. When we hear a salient sound, the eyes will search for the cause of the sound. Attention is brought to the source. We need to select the right object out of a spectrum of objects on the scene. For this reason sounds must create expectancies of what we are about to see. This is possible simply because we know how events and objects sound. And, even more important, the eye is looking for synchronization between the sound and the object. We are, to take one example, sensitive to the synchronization between lip move-

---

[272] North, A.C. & Hargreaves, D.J. (1999) Music and driving game performance. *Scandinavian Journal of Psychology* Vol. 40. No 4. 285–292

[273] Marshall, S. & Cohen, A.J. (1988) Effects of musical soundtracks on attitudes to geometric figures. *Music Perception,* 6. 95–112

ments and speech. It has been shown that babies attend to voices in synch with lip movements and lose interest if the lip movements are unsynchronized. This is in fact exactly what the Marshall-Cohan theory postulates.

It is a nice experience to listen to music as we drive a car. The flow of visual impressions seems to match the music. The music unfolds over time just as the landscape does when we pass through it. At times you can get the impression that the music is especially composed to accompany your trip, especially if there are rhythmic features of repetition and variation in the landscape (telephone poles, house façades etc.). Here is another example that can be explained by the Marshall-Cohan model. If attention selects visual correspondences to the music structure, we would feel that the music fits the visual impression.

The examples demonstrate that place can be perceived in music, but this perception is individual. The visual perception of a location depends on the situation and on the listener. This should rightfully be labelled imagery since there is no visual input. This imagery can be seen as a part of the perception of the music.

Implications for emotion

What is the implication for emotion from this comparison of navigation with music listening? There are several possibilities. I am going to discuss feed back on anticipation (Meyer and Huron), the seeking system (Panksepp), technique, the joy of moving, and flow.

Emotional feed-back on anticipation

Two related theories treat music structure and emotion: Leonard Meyer's theory in *Meaning and emotion in music* from the 50s (Meyer 1956) and David Huron's ITPRA (imagination response, tension response, prediction response, reaction response, appraisal response) theory in *Sweet Anticipation* (Huron 2006). Both deal with expectation. The importance of this is confirmed by physiological studies showing that strong emotional responses to music are typically found in familiar music and in violations of expectancies (Altenmüller 2006)[274].

Meyer's point of departure is Gestalt theory, whereas Huron's approach is perceptual learning. To put it simply, Meyer argues that when the expectations formed by Gestalt principles are fulfilled we are happy and when they are violated or prolonged we are less happy. Are musical "goals", such as the tonic chord in a cadence, really that important to us? Why would we feel disappointed or frustrated if musical goals were not reached? There is no simple reason for this. However ERP research shows, as we have seen, that we do react to violation of expectations.

Huron states that music listening creates expectations and that when these expectations are violated we react negatively. This then would rather explain why we do not like music, than why we like it. Huron offers a solution to this problem. The solution lies in the discrepancy between the immediate pre-attentive reaction and the slower "thoughtful" reaction as we realize that this cannot hurt us. Prefrontal cortex then sends an inhibiting signal. Now we are so relieved that the experience is positive. The bigger the difference the stronger the positive reaction, Huron argues.

The evidence that frontal parts are inhibitory on amygdala reactions can be exemplified by the case of Shepherd Coleman (Tramo 2003)[275]. Mr. Coleman, who is a composer, suffered a rupture of an artery that lies on the surface of the front of the brain. When this artery burst a large amount of blood spilled out and resulted almost immediately in coma. Coleman recovered but there was a dramatic change in his perception of music.

Tramo:

> When he (Coleman) played a piece of music that previously he would find emotional, after the brain damage he was unable to control those emotions and would break out into tears, for example, at a moving part of the piece… There is one part of the opera

[274] Altenmüller, E. (2006) Emotions in music are mainly learned. Paper presented at the conference *Music in the Brain* at the University of Aarhus 21–22 April

[275] Tramo, M. & Coleman, S. (2003) interviewed in Emotion and the Brain in *The Gray Matters* radio series is produced for Public Radio International, in association with the Dana Alliance for Brain Initiatives. www.mega.nu/ampp/danamusic

> (Madama Butterfly of Puccini) where Madama Butterfly sings a duet with her handmaiden, that he simply cannot bear.

Coleman:

> …the duet they sing is of such intensity that I can't even listen to it. I'm discomforted by it. I mean it's painful.

According to Tramo's interpretation of the case, the damaged frontal parts have an inhibiting effect on the limbic system. This is in line with Blood's study on pleasant/unpleasant music, since it was found that pleasant experiences correlated with prefrontal activity. Indeed, activation in prefrontal cortex has been associated specifically with training in the up-regulation and down-regulation of emotional responses (Williams *et al.* 2006[276]; Bonanno *et al.* 2004)[277].

Huron's theory is, as can be seen, grounded in scientific evidence and yet it is unsatisfying. Here are the reasons:

1. Do we ever think that music cannot harm us? If so, we should be aware of such a thought. It never happened to me. The inhibitory effect from the prefrontal cortex does not need to be grounded in reasoning. There is experimental evidence of prefrontal cortex inhibition of amygdala anxiety reactions (Quirk & Milad 2002). The fact that this particular experiment was performed on rats shows that thoughtfulness is not required. Rather the inhibition is due to implicit learning. This is to say that when we have been exposed to the threat several times without bad consequences, the warning is automatically inhibited.

2. Huron's theory would lead to the consequence that the stronger the immediate negative reaction, the happier we are. This is doubtful. My experience is that there is a limit

---

[276] Williams, L., Brown, K., Palmer, D., Liddell, B., Kemp, A., Olivieri, G., Peduto, A. & Gordon, E. (2006) The mellow years? Neural basis of improving emotional stability over age. *The Journal of Neuroscience.* Vol. 26. No 24. 6422–6430

[277] Bonanno, G.A., Papa, A., Lalande, K., Westphal, M. & Coifman, K. (2004) The importance of being flexible: the ability to both enhance and suppress emotional expression predicts long-term adjustment. *Psychological Science.* Vol. 15. No 7. 482–487

for violations. If the violation is found to be within style, it is liked, but if it is out of style, it is disliked.

3. Even if musical surprises evoke emotional reactions, music can be emotional without being surprising. For instance plain rock and roll music can make people go nuts, but there is not much surprise in it. I would say that the music that touches us the most is not seldom quite banal and shallow.

4. The theory is too one-dimensional. The reactions to deviances are positive or negative. Thus the theory does not account for the variety of emotions that music may induce. The most common and the most easily identified musical emotion is sadness (lost love). The theory does not account for sadness.

5. In order to make his theory fit to musical chills (often in combination with piloerection) Huron states that the fast emotional reaction is aggression. His reason for this is that aggressive dogs and cats raise their hair. My own experience is that the chill feeling is a beautiful experience. It is close to love and longing. It seems unlikely that we have to go through aggression to experience this in music. And that this should take place in milliseconds.

In all, I think Huron's theory is interesting, but it needs calibration.

Seeking & emotion

The perception of music and the perception of the landscape resemble each other in a phenomenal as well as in a neurological way; so much in fact that the question must be raised whether the perception of music structure is not an example of cross-domain mapping (see the introduction). This means that our neurological equipment for the basic domain of navigation is used in the more abstract domain of music perception. In this context the research and ideas of Jaak Panksepp becomes interesting.

Panksepp argues that emotions have roles in specific systems (Panksepp 1998). One such system is *the seeking system*. This system is connected with the transmitter substance dopamine. Dopamine is chemically close to cocaine. On a neuronal level the neurotransmitter dopamine is associated with prediction and award (Schultz, Dayan & Montague 1997)[278]. When this drug is injected in mice it typically makes the mice sniff around. This is a very factual example of the close connection between emotion and action.

Dopamine is generally associated with motivation and movement. It has also been associated with the internal clock used to time durations in the striatum (Meck 1996)[279]. Striatum is, as we have already pointed out, involved in the sequence circuit. Panksepp & Bernatzky (2002) speculate that Meck's finding could be connected with "the seeking system". The authors connect this observation with music listening:

> Such basic brain mechanisms for anticipatory eagerness may generate seeking states which may promote various musical expectancies, especially those related to rhythmic movements of the body which may be ancestral pre-adaptations for the emotional components in music. (Panksepp & Bernatzky 2002:136)

This emotion then is connected to curiosity, the eagerness to explore what is hiding behind the next corner. This curiosity of course depends on what we anticipate and the possibility of interesting surprises.

One way to produce musical surprises is to provoke perspective shifts in the listener. The technique of such changes is common. I would recommend listening to Pat Metheny's *The Way Up*[280], which exhibits a map of perspective shifts, all within the allocentric category of perspectives. The perception of the attended part of the music changes because the musical context is changing.

---

278 Schultz, W. Dayan, P. & Montague, P.R. (1997) A neural substrate of prediction and award. *Science*. Vol. 275. 14 March. 593–599

279 Meck, W. H. (1996) Neuropharmacology of timing and time perception. *Cognitive Brain Reseach*. Vol. 3. No 3–4. 227–242

280 Metheny, P. (2005) CD: *The way up,* Nonsuchrecords

1. *Rhythm:* The rhythmic focus in a polyrhythmic beat changes from one rhythm to another. Foreground becomes background and background becomes foreground. Or polymetric effects (as described earlier in this chapter) where the melody does not fit the time signature. It creates rhythmic ambivalence.

2. *Melody:* An intertwining voice in a polyphone weave is suddenly focussed and becomes melody. Or two separately introduced themes are combined to support each other. They give each other new meaning.

3. *Harmony:* A repeated arpeggio indicating a chord is tilted by the introduction of an unexpected base note, giving the arpeggio a new harmonic meaning.

4. *New sonic context:* A sudden simultaneous and total change in sound. You get a new impression of the music. The sensation is so strong you can hardly believe you are listening to the same piece, even if the tonal content is identical.

The effect of these shifts is that you experience a constant re-perception. The music forces you to take new perspectives. It is exactly as in the old pictures used to prove the point of figure and ground in Gestalt psychology. It could also be introducing a new piece of information that produces a change in perception. Metheny's music constantly produces these shifts. This creates awe and you might start imagining your own mind at work. It is like an exploration journey.

The perspective shifts, particularly in rhythm, are bodily felt. Now your feet follow the fast beat and your body sways in a much slower beat layered in the first beat and simultaneously you experience the triplets as waves in your body. As the melody changes focus from one of those rhythms to another there occurs a change in rhythmic focus. These sonic changes produce changes of your body focus. It is not only enriching – it is mind blowing!

Anterior cingulate cortex is involved in a form of attention called attention for action (Posner & Dehaene 1994)[281]. In other words, when surprises occur concerning inappropriate movements, as when we adapt to a new rhythmic pattern, we are alerted. What is the emotional effect? It is totally engaging and you forget everything else. Attention is focussed on to the here and now action. I connect this with the well-known concept *flow* (as described in Chapter 2). Music is one of the activities Csikszentmihaly associates with flow. The self (the conscious thoughts of the self) is forlorn and the subject merges with the task. The experience of merging with the music is commonly reported in the Gabrielsson & Lindström Wik (2003) material on strong experiences related to music.

Technique & emotion

Although it is not difficult to predict music, it is less easy to see what the reward would be. Is the fulfilled expectancy a reward? A pianist was asked to play on a piano where two keys a semitone apart were reversed on the piano (Maidhof 2007)[282]. The pianist showed an ERP component called ERN (error-related negativity). The ERN response has been described as the inappropriateness of self-monitored action within this given context (Luu & Petersen, 2004)[283]. Error, or mismatch, is evaluated in relation to the context of the action. Violations of action relative to action context are essentially expectancy violations. In plain English: you do something and it comes out wrong. I propose, even if this is not spelled out, that the ERN-response in this case could be seen as a measure of technical error. Even if the pianist did the right move, the result was not the expected. The ERN reactions were retrieved both from playing and, although weaker, listening to the error.

[281] Posner, M.I. & Dehaene, S. (1994) Attentional networks. *Trends in Neuroscience.* Vol.17. No 12. 75–79

[282] Maidhof (2007) Neurophysiological correlates of error detection in performing musicians. Poster abstract at the Cognitive Neuroscience Society Annual Meeting Program 2007.

[283] Luu, P., & Pederson, S.M. (2004) The anterior cingulate cortex: regulating actions in context. *Cognitive neuroscience of attention* (Ed. M.I. Posner). Guilford, New York

Why should this experiment be interpreted as a test of technique? The pianist did after all not commit a technical error. The point is that the pianist perceived that she had made an error since the wrong note was heard. There was a mismatch between the performance and the expected feedback.

The ERN on errors in monitored actions in relation to the task are cingulate cortex reactions. The cingulate cortex is generally known to feed back body states, and specifically for signalling if a change of behaviour is required due to deductions of expected outcomes (Bush *et al.* 2002)[284]. The area is activated when ongoing actions are inadequate. The cingulate cortex has, as may be recalled, been pointed out as a part of the singing system. It is activated by listening to music as well as by imagery tasks of music and of imagery tasks of movement.

The anterior cingulate cortex is functionally divided in a cognitive component tracking parameters of the task (context/structure) and an affective component in evaluating response outcomes. It is also known that dopamine is involved in the prediction of future rewards and the motor control needed to achieve the goal. It has been suggested that dopamine is implicated in the striatal involvement in the "*chunking" or automatization of behavioural sequences* (Daw 2003)[285].

The problem with ERPs such as the ERN is that they only detect errors. For this sake I want to direct the reader's attention to an experiment that indicates positive feed back. Unfortunately this experiment did not involve music but it came pretty close. Ullsperger & von Cramon (2003)[286] designed an fMRI experiment where the participants were to foresee which of two balls presented on a screen would hit the wall first. The balls were presented for a sec-

---

284 Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A. & Rosen, B.R. (2002) Dorsal anterior cingulate cortex: a role in award-based decision making. *Proceedings of the National Academy of Sciences*, 99. 523–528

285 Daw, N.D. (2003) Reinforcement learning models of the dopamine system and their behavioral implications www.gatsby.ucl.ac.uk/~daw/thesis.pdf

286 Ullsperger, M. & von Cramon, D. Y. (2003) Error monitoring using external feedback: Specific roles of the habenular complex, the reward System, and the cingulate motor area revealed by functional magnetic resonance imaging. *The Journal of Neuroscience*. Vol. 23 No 10. 4308–4314

ond. A correct answer (pressing the right button) was rewarded with a smiling face and vice versa. The resemblance with playing an instrument is this:

1. We have the perception of movement related to a landmark in both cases (tonal centre and a wall respectively).
2. Both cases involve a prediction of movement that is procedural rather than calculated.
3. Both involve button or key pressing as responses to prediction of movements (seen or heard).
4. There is a feedback in both cases – a smiling face and the sound of the right tone.

Ullsperger and von Cramon found that positive feedback from the button pressing correlated with activity in the ventral striatum. This area, also known as the nucleus accumbens is activated by pleasurable experiences. It is an attention area, rich in opioids such as endorphin receptors and considered implicated in drug addictive behaviour. Negative feedback was correlated with activity in the rostral cingulate motor area, the inferior anterior insula and the epithalamus (habenular complex). The cingulate area is as we have seen generating ERN error related negativity and feedback generated negativity. The epithalamus is known to have inhibitory effect on midbrain nuclei (such as nucleus accumbens) and thus reduces levels of dopamine.

Ullsperger and von Cramon's finding about the neurological underpinning of positive emotions in this task is in agreement with an investigation of musical emotions (Menon & Levitin 2005)[287]. Menon and Levitin found that music caused a correlated activity in the ventral striatum and the nucleus accumbens. It could thus be argued that we do not only react to fulfilled expectancies and violations in music but also to performance outcomes. Thus, if we latently take part of the performance as listeners, there is a possibility that we would feel

[287] Menon, V. & Levitin, D.J. (2005) The rewards of music listening: Response an physiological connectivity of the mesolimbic system. *NeuroImage.* Vol. 28. No 1. 175–184

the lack of technique in our own bodies. And on the contrary, the perception of perfect technique would be perceived as pleasant – "beautiful". In fact, this is two sides of the same coin. The perfect technique is what we learn to expect and the false tone is a violation of this expected perfection.

To sum up:

1. Technical errors cause ERN reactions in the anterior cingulate cortex.
2. Positive feedback on appropriate action monitoring activates the ventral striatum and nucleus accumbens.
3. The listener feels the perfection of technique in her own body.

I think that no theory of music and emotion should ignore that there is a world of a difference between listening to tight music on pitch performed with ease and listening to the opposite. It is simply the difference between listening to the skilled professional and the amateur.

I propose that the perception of technical excellence is an important explanation to positive emotions in music. Basically it is the same bodily satisfaction as when we manage to hit a target perfectly. It is the sensation of a perfect backhand volley. Or seeing someone else perform it. It is the feedback of success telling our bodies: This is how to do it! We feel the elegance of the move. Just as much as we love seeing Ronaldinho make a move, we love the perfection of Francisco Sánchez Gómez, alias Paco de Lucia. And just as we struggle with the wrestler fighting to avoid losing by a fall, we struggle when we listen to out of pitch singing.

The joy of movement

Even if technique is fun, let us not forget that movement is fun generally. I will not expand on this since we can all relate to the truth in this from our own experience. Just let me point out one interesting observation! Striatum organizes sequential movements. In this process it releases serotonin (Jacobs & Fornal 1997)[288]. This could explain such diverse activities as jogging, rocking babies to sleep and gum chewing.

From a musician's point of view the striatum-serotonin reaction on repetitive action, as well as the striatum-dopamine reaction on timing, is highly interesting. Music consists of repetitive patterns which in turn activate repetitive body action. This might explain why we enjoy endless rhythmic repetition. The extreme cases are rock music with its monotone pulse, techno which is nothing much but rhythmic sequences, rave dancing, the festival in Rio de Janeiro, where you choose one rhythm in the rhythmical weave and beat it enthusiastically for a whole week, or the continuous monotone rhythms in shaman music, which put the listener in alternative states of consciousness. Since even the so called classical music is made up by rhythmic and melodic repetitive patterns and since there are always motor tendencies in listeners this effect can always be assumed to some degree.

Conclusion

We have discussed several phenomenal and neurological parallels between music listening and navigation. This connection indicates *cross-domain mapping* defined as a general cognitive process through which we structure an complex or abstract domain (music) in terms of one more familiar or concrete (finding the way in a familiar surrounding). This has many implications for music and emotion: feed back on anticipation, the seeking system, technique, the enjoyment of movement and flow have been discussed.

Allocentric emotions, if I may call them so, reflect the enjoyment and frustrations of following a path; our excitement of what might be around the corner, reunions and surprises, as well as the thrill to master this technically. All this contributes to the flow, the alert sense of here and now. This makes us merge with the music. Let me accentuate that this here and now experience, this merging ex-

[288] Jacobs, B.L. & Fornal, C.A. (1997) Serotonin and motor activity. *Current Opinion in Neurobiology*. Vol. 7. No 6. 820–825

perience, is exactly the characteristic of the strong experience in music.

ERP experiments show that this emotion does not result from what we consciously perceive. Rather emotion is generated earlier in the perception process. We recognize this from the egocentric and the dyadic perspectives.

5

Tribal Rites

The Tribal Perspective

Music has two intriguing effects:

1. It coordinates people physically.
2. It makes us feel the emotion of others. It creates social bonding and may therefore be essential for maintaining large groups. This, David Huron argues, can be the evolutionary explanation why we have developed music (Huron 2003). The large group had an advantage in the feuds with other groups and in detecting, hunting and conquering big animals. It can thus be argued that music is beneficial for the spread of the genes of the tribe.

I attempt to show that coordination and bonding are connected and that this connection creates tribal cosmology – a shared perception of the world.

Depiction of perspective

As can be seen from mass psychology reactions such as panic and riot rage, it is possible to perceive the intentions, the goals and the emotions of the group, not as a bunch of individuals, but as a unit. I will call such a unit a *tribe* – a group involved in, and with a history of, bodily interaction. Interaction with the tribe creates *the tribal perspective*. The frame of reference is the information shared by the tribe. We are as listeners affected by performing groups, ensembles, and orchestras; but also by the audience. Audiences are shaped to tribes by the performers, as are music groups by their audiences. They belong to the same tribe.

Tribalism can be felt at live concerts, carnivals, football crowds, dance-floors and rituals. When we take this perspective our roles, values and lifestyles are dictated by the tribe. This can be more or less pronounced. It is evident in teenage audiences, but even the bourgeois elderly couple attending classical concerts belong to a tribe, although to a less pronounced degree. The music and the audience affect this couple, change their behaviour, their movement pattern, and their perception of who they are.

Music is a sensitive subject, because it shapes identities. We internalize the features of Metallica, say, and make the emblems, the tattoos, the clothing, the movement pattern and the attitude our own. If someone ridicules your music, you – as a person – are under attack.

What is the underlying mechanism to tribalism in music? I will introduce and compare evidence and theories about group interaction and shared perception. I will then discuss three examples: The carnival in Rio de Janeiro, the song Sangrando by Gonzaguinha and the rock concert.

Tribal life

In the dawn of manhood semi-nomadic hunter-gatherer groups of 10–30 people lived face-to-face with each other in extended families (Oatley & Jenkins 1996)[289]. The total number of acquaintances, including other groups, has been estimated to about 150. This is the life *Homo Sapiens* (as Carl Linneus classified humanity) was adapted to live and it is the life we have been living during the predominating part of our history. We are talking about the pre-cultural era

[289] Oatley, K. & Jenkins, J. (1996) *Understanding Emotions*. Blackwell, Cambridge, Massachusetts

starting approximately 160.000 years ago and lasting until 80–40.000 years ago when the first cultural artefacts emerge. The reconstruction of this early era is a key to the understanding of who we are. Advanced guesses on tribal life can be made from paleontological research, and estimations of resources, tasks, problems and threats that had to be dealt with to adapt to the environment. Before language was fully developed communication was restricted to bodily interaction, to concrete showing, to gestures, to the tone of voice and to spontaneous outbursts. What had to be communicated in tribal life? And how is this reflected in our behaviour today?

The hierarchy within the group had to be fixed. Peace within a group is maintained by an acceptance of the hierarchy to avoid internal fights over food and mating partners. This means that dominance and submission had to be communicated. The group decides who you are, not only in the eyes of the group but in your own eyes. This communication can also be seen in herd living mammals.

In the tribal perspective you can have identity (a role in the group) but you cannot have individuality. We can perceive the self as a co-perception of the group, but individuality demands cognitive consciousness. The thought "This is me" must be formulated. A life story must be constructed. This probably demands language. If so, it is a cultural perspective. These two perspectives can come into conflict. Let me exemplify!

At school reunions, even the most successful person can be reduced to the little scared and bullied child the class once doomed her to be. This decision is hard to change. The exposed girl loses self-confidence, the voice becomes weak, she not only experiences that the group thinks that she does not have anything of interest to say, but even feels that this is true. Everything comes out wrong. She confirms the expectations. She may have achieved many things in life, but this is worth nothing when she meets with the class.

In cases like this, it does not matter that we are cognitively conscious of our identity. The phenomenal consciousness of the group dominates over the cognitive consciousness of individuality when we interact with the group. Amazingly, we tend to accept the role ascribed to us. It is the condition to belong to the group.

It could be argued that this depends to some extent on linguistic communication. But body language is enough. How does the group look at you when you speak? Do you get attention? How do they address you in their tone of voice, gaze and body language? Do they abuse you or threaten you? Bullying is hard to change with logical arguments because it is not logically motivated.

Thus in one group you may be held in high esteem and in another you are pariah. You may be a hero at work, but despised at home. Every group, from the day care centre to the geriatric care centre, is a tribal world with values of its own.

An interesting adaptation to the group, which can be seen also in animals, is the sacrifice for the group. Some group members may risk their life for the survival of the group (Matura & Varela 1987)[290]. The tribal perspective thus can dominate over the egocentric perspective, even if it means danger to the individual. The struggle for survival is not just egocentric but also tribal. This reflects that our genes stand a better chance if we protect the group (containing our offspring) than if we protect ourselves.

Here we can see a parallel in traditional military training which aims to erase critical thinking. Soldiers are not taught to sacrifice their lives by verbal education but through enacting military drill, which is typically a matter of synchronizing the group physically to one movement and one will.

Such preparation was essential in the tribal communities. If the warriors were not prepared to sacrifice their lives, the whole tribe would die, or, in the case the women were spared, they would not be available to the warriors again.

Roles are given in daily interaction. Here we can see the correspondence between roles and emotions. Attachment, care giving, cooperation, competition, and predation are tested in recurring situations as joining (happiness, love), interruption of attachment (distress and anxiety), resumption of attachment (relief), loss of attachment (despair), assisting others (love), formation of relations (happiness), exchange (gratitude), intercourse (sexual feelings), loss of relationship (sadness), achieving and defending resources (anger), defeat

[290] Maturana, H. & Varela, F.J. (1987) *The Tree of Knowledge: The biological roots of human understanding*. Shambala, Boston

(shame, revenge), victory (pride), escape (fear) (Oatley & Jenkins 1996). We call this special category of emotions *relational emotions.* They can emanate both from dyadic and tribal perspectives.

Roles are also given in rituals. In baptism, initiation rites, marriages, and funerals members are guided by the shaman, the medicine man, or the priest through a transition to a new role. The shaman has a mediating role between this world and the spiritual world (the myth). The passage to new consciousness is often a state of trance induced by music.

Trance

The aim of the rite is often some kind of passage. This passage can be spiritual journeys or the transition of identity as in the Afro Brazilian religions where participants dance themselves into trance to be possessed by *orixas* (gods).

Trance is an *altered state of consciousness.* The ritual (including music, dance and drugs) is a technique to reach such a state. Other altered states are achieved by hypnosis, meditation and flow. Common to these states are a lowered sense of cognitive consciousness and a heightened phenomenal consciousness resulting in unreflective presence. A common factor in all meditation techniques is the focusing on something simple and concrete such as breathing, a sound, a mantra, a body part, the pendulum of the hypnotizer.

Thus altered states of consciousness are achieved by focussing on something present. Thinking is repressed. Thinking is an indirect experience of the world where the world is processed with symbols. In trance there is an absence of inner language (Friedson 1996)[291]. In the state of trance cognitive consciousness is not possible. The altered state is a phenomenal perception of the world. As a consequence critical defence breaks down. Suggestibility increases. This is a world where episodic memories and imagery is lived and experienced. It is a dreamlike state although the participants are fully awake. The heightened suggestibility makes the trance experi-

[291] Friedson, S. (1996) *Dancing Prophets: Musical experience in Tumbuka healing.* Chicago University Press, Chicago

ence a passage. You come out on the other side with a new identity.

The Indian Tantric philosopher Abhinavagupta described how consciousness changes from an individualistic state of "contraction" to a ritual state of "expansion":

> The consciousness, which consists of, and is animated by, all things on account of the difference of bodies, enters a state of contraction. But, in public celebrations, it returns to a state of expansion – since all the complements are reflected in each other. The flow of one's own consciousness in ebullition (i.e., when it is tending to come out of itself) is reflected in the consciousness of all the bystanders, as if in so many mirrors, and, inflamed by these, it abandons without effort its state of individual contraction. (Abhinavagupta in Gnoli 1956: 70)[292]

This text from the 11th century gives an amazingly precise depiction of the tribal perspective. It describes how the own body, as it is reflected by the bystanders in the celebration, loses individuality. The partakers cognitively consciousness of themselves as individuals (cultural perspective) is replaced by phenomenal consciousness (tribal perspective). This is felt as an expansion (and a return!).

Music is, as we have seen in the preceding chapters, highly entrainable because of its predictability. The listener becomes part of a collective movement. We do this pre-attentively, but we can also as active listeners entrain it attentively – get absorbed by it; cut out everything else, and thus reach a stage of altered consciousness.

Judith Becker (2004)[293] compares reported experiences of trance with reports of *deep listening* to music. Trancers, like music listeners, follow a "script". Common experiences are: a loss of the strong sense of self, merging (with an object such as music or with others), changed perception of space and time, a kind of syncope – a real bodily change, such as fainting, experienced as a blessing. Another common thing is that the experience as such is difficult to describe.

---

[292] Gnoli, R. (1956) *The Aesthetic Experience According to Abhinava Gupta.*: Chowkhamba Publications, Varanasi, India

[293] Becker, J. (2004) *Deep Listeners – music emotion and trancing.* Indiana University Press, Indianapolis

Becker reports *trance amnesia* – inability to recall what transpired during trancing. Something like this may happen in strong music experiences:

> I find it very difficult to find words to this music that I have experienced so strongly at one occasion. The closest description I can come up with is that it was a cosmic wholeness experience beyond time and space. The body and the music became a whole… (Gabrielsson & Lindström 1993)[294]

This trance amnesia may have to do with the lack of inner language during the experience. When the experience is verbalized and thus transposed from one perspective to another, there is an obvious risk for rationalizations. The experience, if recalled at all, will be depicted and interpreted differently in different cultures. Additionally, the lack of inner language could explain the reported loss of self. If we cannot symbolize, we cannot formulate who we are. The so-called "autobiographical self" vanishes. The autobiographical self is a semiotic reconstruction of episodic memories to a narrative. It is a rationalization – an effort to make our lives understandable.

Trance is the objective in *rave* dancing, particularly in the style called *Goa Trance*. Rave is a *neotribalistic* phenomenon.

> The dance space in trance-dance parties is a sacred space. It is a form of meditative, collective, spiritual worship. It is a reconnection with the elemental primordial rhythms of organic, cosmic life force. This sonic, mantric transcendence state is heightened when people dance together in nature under the stars and experience the celestial shift of dark into light with sunrise (…). There is a transmission of higher levels of knowledge through the song and dance, which is a direct conduit for the religious experience and unites them tribally (…). (The duo Insectoid 1997)[295]

Goa Trance music is minimalist development over 8–9 minutes. Over a 4/4 rhythm, with a base pumping 16th bourdon notes strong

---

[294] Gabrielsson, A. & Lindstöm, S. (1993) On strong experiences in music. *Musik Psykologie.* 10. Florian Noetzel Verlag, Wilhelmshafen. 118–40

[295] Insectoid (1997) quoted from the folder of the CD *Sacred Sites.* Produced by Return to the Source

enough to rearrange your bone structure, you will hear short repetitive melodies (if any at all). Synthesizer futuristic space spiced with sampled whales or birds to signal the Gaia (Earth Goddess) connection, alternatively sampled sounds from Tibetan monks or Hindi processions to signal archaic non Western religion, are interspersed with mantras as "You can be Shiva, you can be Shiva".

This music, which started in Goa India and Ibiza in the 80s with British disc jockeys as priests, to explode at warehouse parties and outdoor events in the 90s, is perhaps a too obvious example of modern tribalism. However, even if the discourse has changed, the influence on new dance styles is strong. It works as a mean to achieve trancelike states on the dance floor, particularly in combinations with drugs.

The crowd

Gustave Le Bon, a pioneer in mass psychology argued in 1895 that individuals in a crowd, irrespective of individual background, just by their transformation to a crowd, develop a "group-soul" – a creature with radically other features than the including cells (Le Bon 1895)[296]. Here, Le Bon stated, appears *the species specific unconscious.*

The subject has interested intellectuals as Sigmund Freud, Mikhail Bakhtin, Elias Canetti and Daniel Stern – Freud because of his interest in the subconscious, Bakhtin in his depiction of the carnival, Canetti because of the link to power and Stern in his investigations of the present moment.

The crowd releases, according to Freud, unconscious impulses from repression (Freud 2004)[297]. Each emotion and action becomes contagious. There is a likeness, he noted, between what takes place in crowds and what happens in hypnoses – a suggestibility. A compulsion is developed to harmonize with the group and a pleasure to let one self go, to melt into the crowd completely. In crowds we regress to a lower level – an earlier level, where

[296] Le Bon, G. (1982) *The Crowd: a Study of the Popular Mind.* Atlanta: Cherokee Publishing Company

[297] Freud, S. (2004) *Mass Psychology and other Writings.* Penguin, London

regress to a lower level – an earlier level, where emotional life is stronger and critical consciousness declines.

Here we can clearly see a parallel to Trevarthen's description of protoconversation (Trevarthen 1999). In the infant's dialog with her mother we see the mechanisms of entrainment, perceived expectations, and the rhythmic communication. This likeness is not coincidental. The common factor is imitation. In protoconversation as in crowds we are attuned to the surrounding.

Bakhtin identified a moment when a group of people changes from being a composite of hesitating individuals to become one gestalt animated by one spirit (Bakhtin 1984)[298].

Canetti stated that the most important occurrence within the crowd is the discharge. Before this the crowd does not actually exist; it is the discharge, which creates it. This is the moment when all who belong to the crowd get rid of their differences and feel equal (Canetti 1993)[299].

In the interaction with other people's feelings and thoughts, what is ours and what belongs to other people starts to break down (Daniel Stern 2004).

What is happening in these transformations? Ever since the knowledge of mass effects became systematically used in Nazi propaganda, this line of research has been avoided. There is however in biology and in cognitive science an awakening interest of how we affect each other in group interaction. I will briefly present two directions: *structural coupling* and *joint-action* theory.

Interaction-interperception

Interaction causes change in the participants. This can be noticed on any biological level. The phenomenon has been labelled *structural coupling*.

---

[298] Bakhtin, M. (1984) *Rabelais and His World.* Indiana University Press, Bloomington

[299] Canetti, E. (1993) *Crowds and Power.* The Continuum Publishing Corporation, New York

> We are speaking of structural coupling whenever there is a history of recurrent interactions leading to the structural congruence between two or more systems. (Matura & Varela 1987: 75)

Structural coupling is a long-lasting biological internal structural change in the individual as a result of a common history of group interaction. These changes are described as knowledge, gained through collective doing. Knowledge is thus understood as a *structural change* resulting from interaction, not more. Communication is described as "coordinated behaviours mutually triggered among the members of a social unity" (Matura & Varela 1987: 193).

The *structural congruence* is an adaptation to the other(s). It creates a shared context, in an elementary sense – a shared perception of the world. Matura & Varela's book *The Tree of Knowledge* is one long argument for the thesis that the only world we can have together is the one we create through the actions of our co-existence.

The depiction of structural coupling makes it a perception-action theory. Action leads to perception. If we have an interacting group, interaction will lead to *inter*perception – a tribal cosmology. Let me take an example.

Attention weighting refers to our ability to unconsciously direct our attention to features that are important to us. This is an example of *perceptual learning*. In tribal interaction some features become important and are thus better perceived than others. For teenagers music style is an important feature. Through daily interaction with the music the rhythms and the moves are internalized. Music style is linked to tags, clothes, movement patterns and attitudes connected with tribal identification. Music style becomes a very nuanced and rich world. Style features are spotted immediately. Wrong music – wrong tribe.

Group interaction is focussed by a new branch of cognitive science called *joint action* (see Sebanz, Bekkering & Knoblich 2006 for overview)[300]. It concerns the ability to coordinate our actions with those of others. Joint action is, among other things, a state of devel-

---

[300] Sebanz, N., Bekkering, H. & Knoblich, G. (2006) Joint action: bodies and minds moving together. *Trends in Cognitive Science.* Vol. 10 No. 2. 70–76

opment in children. Playing games with others starts at 12 to 18 months of age.

It is sometimes said that the ability to coordinate with others requires a “theory of mind” – a hypothesis of the state and the intention of the other. This is however not necessary since joint action is considered to be implicit. It may sound amazing that such complicated behaviour could be performed without cognition. But, as we all know many species act jointly. We find joint action on the level of infants, mammals (hunting herds of wolves), birds (flocks in plough formations), fish (shoals) and even insects (swarming). It is a widespread biological phenomenon. And it is poorly understood. Is there a common mechanism behind these interactions?

It has been suggested that joint action can be explained by mirror neuron theory (Rizzolatti & Craighero 2004). By watching the action of the other(s), exactly the same action is simulated in the observer. The simulation is embodied since it uses a pre-existing brain model of the body – *the body schema* (see Chapter 2). For this reason, the term *motor resonance* (Wilson & Knoblich 2005) is used.

As discussed in Chapter 2 mirror neuron theory may explain not just imitation, but also complementary action, intention understanding, situation understanding including goals, prediction of movement of the other, entrainment, empathy and complementary emotion. This is important for joint action which is not just about coordination in unison, but also, as in the case of herd hunting and children’s games, about the ability to make complementary moves. This demands the ability to sense the goal of others, to predict actions, to integrate predicted effects of own and others actions and – *the ability to share representations*. Joint action can thus be specified as interaction depending on:

1. Shared perspective. The “perceptual common ground” (Bayliss & Tipper 2005)[301] is the prerequisite as well as the result of joint action. Such a common ground is the sharing

[301] Bayliss, A.R. and Tipper, S.P. (2005) Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology*. Vol. 96. No 1. 95–114

of time, sequences and rules. This entails shared predictions of movements.

2. Common action goals (Umiltà *et al.* 2001; Kohler, E. *et al.* 2002[302]; Fogassi *et al.* 2005[303]). This entails shared imaginations of outcomes.

3. A sense of the intention of the others.

Observing errors committed by others furthermore activates the same neural frontal activity as own errors in similar tasks (van Schie *et al.* 2004[304]). The mere observation of erroneous action leads to pre-motor activation (Manthey *et al.* 2003)[305].

Motor resonance, complementary action, predictability, shared perspective, shared goals, and the sense of the others intentions and emotions create a sense of who contributes with what – the distribution of work – to achieve a common goal. Interestingly the feedback from the timing of the other's actions can become as effective for anticipatory action control as internal signals about one's own actions (Knoblich & Jordan 2003)[306]. As a result each member can learn to predict the timing of the other. By sensing the intentions of the others, we can in a second step implicitly understand the situation and know how to contribute. Every member steps into a shared reality depending on the shared perspective. The finding that mirror neurons are active at complementary action (Newman-Norlund *et al.* 2007) throws new light over ensemble playing and improvising. We

---

[302] Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V. & Rizzolatti, G. (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science.* Vol. 297. No 5582. 846–848

[303] Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. (2005) Parietal lobe: from action organization to intention understanding. *Science.* Vol. 308. No 5722. 662–667

[304] van Schie, H.T., Mars, R.B., Coles, M.G.H. & Bekkering, H. (2004), Modulation of activity in medial frontal and motor cortices during error observation, *Nature Neuroscience.* Vol. 7. No 5. 549–554

[305] Manthey, S., Schubotz, R.I., & von Cramon, D. Y. (2003). Pre-motor cortex in observing erroneous action: An fMRI study. *Cognitive Brain Research.* Vol. 15. No 3. 296–307

[306] Knoblich, G. & Jordan J.S. (2003) Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology. Learning Memory and Cognition.* Vol. 29. No 5. 1006–1016

have an automatic urge to fill in and complete what the others are doing. The impulse: if you do that, I will contribute with this, is strong.

We empathically sense the goal and the intention of other people in our own body. According to Metzinger & Gallese (2003) "this functional mechanism seems to play a major role in our epistemic approach to the world". It creates, they claim, a *shared action ontology*. I suggest that the creation of shared action ontology follows these four steps:

1. A mirror neuron induced imitation or complementary action.
2. Joint action creates phenomenal situation understanding.
3. This situation understanding is stored in episodic memory.
4. Episodic memories from group interactions make up shared action ontology.

This deepens our understanding of the tribal perspective as a phenomenal perception of the world through a filter of shared episodic memories.

The discussion of structural coupling and joint action actualizes the definition of communication. As we generalize the meaning of this word to pre-attentive and biological levels the meaning expands from type of transfer (Shannon & Weaver 1949)[307] to type of sharing (Allwood 2000)[308] or just resonance (St Clair & Busch 2000)[309]. As we have seen the type of communication we have been discussing does not just result in sharing in an immediate sense. Complementary action and complementary emotion can also be caused by communication. In this case it could be argued that the perspective is communicated and shared. Tribal communication may however

[307] Shannon, C.E. & Weaver, W. (1949) *A Mathematical Model of Communication.* University of Illinois Press, Urbana, IL

[308] Allwood, J. (2000) Activity based pragmatics. *Abduction, Belief and Context in Dialogue: studies in computational pragmatics* (Eds. H. Bunt & B. Black) John Benjamins, Amsterdam. 47–80

[309] St. Clair, R.N. & Bush, J.A. (2001) Transmission of values: the information age crisis in socialization. *Intercultural Communication Studies XI.* 67–69

result in differences – in hierarchies. And as we all know from conflicts, communication sometimes just accentuates difference. So what is shared in such cases? The fact that we disagree? Perhaps we have to admit that communication can produce sharing as well as difference. If so we might just have to settle with the understanding *production of change.* As soon as something has caused change in something, there has been some kind of communication. In order to come up with more specific definitions we must specify the type of communication.

Are there any indications of this perspective in neuromusicology? New fMRI evidence shows that the planum temporale is engaged by collective speech and singing but not at solo performances (Saito *et al.* 2006)[310]. This area is activated for listening as well. The authors draw the conclusion that this area synchronizes the singer/speaker with the other(s).

Mythology

Let us compare tribal cosmology with the anthropological concept mythology. Claude Lévi-Strauss was captured by the decorations on Brazilian Indian craft works. The figures were interpreted as mythological conceptions of the world (Lévi-Strauss 1955)[311]. Lévi-Strauss developed this theme in *Anthropologie Structurale* (1958)[312] and in *Mythologiques 1–4* (1968)[313]. I am in the following referring to the general message in these works.

Myths to Strauss were structures rather than narratives. The same myth could thus be told in different ways. It is like chess. Every game is a sequence, but under this sequence there is a rule-governed structure allowing for endless combinations. And as in chess there are figures in myths with roles such as the father (mon-

---

310 Saito, Y., Ishii, K., Yagi, K., Tatsumi, I. & Mizusawa, H. (2006) Cerebral networks for spontaneous and synchronized singing and speaking. *Cognitive Neuroscience and Neuropsychology.* Vol.17. No 18. 1893–1897

311 Lévi-Strauss, C. (1955) *Tristes Tropiques.* Plon, Paris

312 Lévi-Strauss, C. (1958) *Anthropogie Structurale.* Plon, Paris

313 Lévi-Strauss, C. (1968) *Mythologiques.* Plon, Paris

ster), the son (wounded hero), and the mediating trickster. The figures act out themes that recur frequently – themes where contrasts come into play. The roles are stereotypes, not individual. We can see such stereotyped roles in rituals, carnivals, in commedia dell'arte, in the puppet theatre and at the circus (e.g. the clowns) – all art forms where half human and half animal characters are played. Each figure wears a specific mask, a specific dress, and has a specific movement pattern.

To Lévi-Strauss the myth was a collectively shared structure. And this structure is the implicit knowledge of the world. Such inner structures are manifested as images in dreams, plays, art, craft works, and rituals. It is often said that these manifestations are symbols. If so, they are symbols in the Freudian-Jungian tradition. The images are however not symbols in the Peircian sense, but icons. The representations cannot be adequately expressed in words. The meaning of the mythological image is darkly divined and beyond grasp.

Structuralists claim that although mythological manifestations of roles and relations of a tribe are specific, there is a universal structure underneath. The structure of myth is isomorphic to the structure of certain mind processes common to all human beings.

As may be recalled from Chapter 2, Merleau-Ponty saw myth not as a representation but as genuine presence:

> Myth beholds the essence within the appearance; the mythical phenomenon is not a representation, but a genuine presence. The daemon of rain is present in each drop…the soul present in each part of the body. Every apparition is in this case an incarnation. (Merleau-Ponty 1996: 290)

This perception-action understanding of myth, and it is the understanding I will use, makes myth almost synonymous with tribal cosmology.

According to this quote the perception of the myth is an incarnation. The myth is perceived as real – a concrete perception of the world. An example of an incarnation of the myth is the hero. The hero is a figure charged with needs, virtues, qualities and expectations that emanates from the people. The hero is often wounded, has shortcomings, doubts, as do we all. But the hero overcomes these difficulties to win a final victory, as we wish we could. This way the hero is the people, inseparably connected to the people. As the Brazilian anthropologist Gilberto Freyre once said, there is in the cult of the hero a touch of the affection of the cat: when seeming to be caressing the leg of the master, what the cat is doing is voluptuously stroking its own fur.

The step from the mythological hero to divinity is not far. Jesus on the cross, wounded, doubting and weak is very human. This is the iconic image of Jesus. This is the image carried through the streets in Catholic processions. It is illuminating that scholars of the new field cognitive anthropology exploit episodic memory in their understanding of how religious knowledge is communicated. It is thus argued that an imagistic mode of religiosity is created by infrequent but emotionally arousing ritual performance (Whitehouse 2000)[314].

Let me take a musical example. Rumour says that Quincy Troupe did not once interview Miles Davies when he wrote *Miles – the autobiography*. He wrote the book from other available sources and when he asked Miles if he had got it right, Miles' comment was: "There is not a word of truth in this book. But it's cool. Print the shit."

In fact the book gives a contradictory impression. On the front page of the book under the picture of Miles in a black coat and a black hat holding his trumpet, you can read: "Miles Davies with Quincy Troupe". But among the long list of contributors in the acknowledgement, Miles' name is not mentioned (Davies & Troupe 1990)[315]. It does not need to be of course if Miles is really one of the authors. But if he is one of the authors, why then did they have to tape hundreds of testimonies? There is thus a possibility that the book is a distillate of information shared among musicians and that the musicians unintentionally have added pieces of themselves and each other to the story. Such myths emanate in musician interaction – in playing, in touring buses, in the jargon. Jazz musicians have a way of talking that seems to be a prolongation of musical group improvisation.

---

[314] Whitehouse, H. (2000) *Arguments and Icons: divergent modes of religiosity*. Oxford University Press, Oxford

[315] Davies, M. & Troupe, Q. (1990) *Miles – the autobiography*. Pan Books, London.

This book has contributed to the spread of the legend. It is not longer possible to listen to a solo by Miles without perceiving the image of the incarnated myth. The icon is present in the tone vibrating of the struggle oscillating between cool artistry and drug destruction, between heavenly beauty and pimp jargon from the streets of Harlem, between good and evil, between life and death. The tone carries the myth.

Imagine Miles music without the myth! Had the sound of the trumpet been as captivating if we thought that the soloist was some ordinary guy – married with two kids, a house in the suburbs, a Volvo and a golden retriever?

This example shows that the myth may travel through different stages. It emanates from interaction (tribal); then it becomes a narration (cultural) – a legend, and finally it is evoked by music listening.

Great musicians tend to be heroes – canonized as geniuses; worshiped as idols. We cannot listen to the music without hearing the hero.

The musician can have another mythological role – the role of the mediator – the trickster, the shaman, the priest moving between worlds. Half human, half spiritual the mediator reveals the other side to the tribe. He invites the tribe to travel in the spiritual world. The technique to make this happen is the ritual.

Rites

To the early tribes human communication was probably largely restricted to indexical pointing, emotional contagion and showing how. In showing how - a situation is mimicked. The tribe enacts the situation and thus the situation is internalized as procedural knowledge. When this becomes a tradition, we have a ritual. The rite is a concrete attempt to make things appear. It is at the same time a collective enactment of a situation and a way to make this situation real. This makes the ritual an education process, a sharing of knowledge in the tribe. The rite produces, communicates and reinforces tribal cosmology. It is a staged drama concerning the identity of the tribe, its members and the conditions of life.

The rite is a representation, but not quite, because the act is actually performed. It is the fascinating border between the real and the image of the real; between the real and communication, and between life and art.

Child games are ritualistic in this sense. What are these games, but an interaction creating a consistent world? Here we detect the stereotyped characters, the fixed roles typical to myths: mummy-daddy-child, chaser-chased, wolves and fairies, witches and goblins.

The Brazilian anthropologist Roberto DaMatta (1991)[316] characterizes ritual is as something displaced and highlighted. Something is taken out of its context, replaced, dramatized and staged.

An important feature of the ritual is that it follows a schema. It is a replication of *joint-action* sequences. The order of events, movements, gestures, and words is fixed. Every event is anticipated and this is why the tribe can entrain the action.

In this sense music is not only a constituent of rituals; it is itself ritualistic. Every time we sing or play a song we reproduce an earlier occasion. Even as we listen, we enact music more or less openly. Entrainment is, says anthropologist Judith Becker, the strongest form of interaction. Bodies and brains synchronize gestures, muscle actions, breathing, and brainwaves while enveloped in collective rhythmic action (Becker 2001). It actualizes, she states, the *supra-individual* state – a collective embodied understanding of the action.

In this respect communal singing is a ritual. Communal singing is heard around camp fires, in churches, in schools, in football arenas, political meetings, and at national celebrations. Just by this brief look at the occasions, we get an inkling of possible functions or effects. Social bonding is an important ingredient in all of them, as is the emotional understanding of who we are and how we are related to common roots and a common cause.

There is, to take an example, no midsummer eve in Sweden until we start singing the old songs that go with the ritualistic cheering with aquavit. It is the singing – that makes the event happen. We

---

[316] DaMatta, R. (1991) *Carnivals, Rogues, and Heroes – an interpretation of the Brazilian dilemma*. University of Notre Dam Press, Indiana

may know it is midsummer, but it is not until we sing the songs that it really happens.

By the same token the ritual is the only way to conjure up religious reality – to reincarnate the words, to really make the wonder happen. We sing in unison in the chorales. We read the prayers in unison. The priest is performing the very moves and speaking the very words priests have uttered again and again for ages. If anything is changed or modernized, as they sometimes try, the magic disappears. For example, we are touched at a funeral when the priest take the spade and spreads earth over the coffin and says: "Ashes to ashes, dust to dust" because, as we follow him, we entrain the exact ritual and the exact text that unites us with the ancestors before us, mourning their deceased fathers and mothers in an eternal chain. It triggers episodic memories of earlier ceremonies – including the emotions. It opens up a bodily connection to the past. The past is reshaped. It happens again. The former funerals happen in this funeral. In a flash we might sense in the wind the universal condition of life – death. It is, for many of us, not until the moment of the ritual, we actually experience the loss. We need the ritual to finally accept the fact, because without the ritual we never really parted.

The perception of the real event, the perception of the myth and the perception of the symbol can be seen as developmental stages in the child as well as in human history. It has an obvious affinity with thought figures in the 20th century: Freud's id-ego-superego, Peirce' index-icon-symbol and Lacan's real-imagery-symbolic. These are, I would say, perspectives.

At first ritualistic communication was perhaps unintentional. But at some point gestures and sounds lead to iconic communication. The gesture, the dance move, the facial expression, amulets, totems, face paintings, and masks and the picture became charged with tribal meaning, as are today clothing, moving patterns, hand signs, tattoos, tags and graffiti – and music sounds.

In the ritual, information is transferred from a source domain to a target domain. When the experiences are acted out collectively, every member or the tribe entrains the experience. Practical life is the source; the myth is the target. This myth is the cosmology of the tribe.

Carnival

In Rio de Janeiro, as in the rest of the world, samba music and carnival are closely associated. I want to stress this because in the following I will discuss how tribal mechanisms colour the perception of music. This could be argued for any music except purely intellectual music constructions. Behind jazz, blues, kletzmer, tango, flamenco, raga, religious music, Arabian music etc. there are tribal traditions important for the perception of the music. Even classical music can be traced back to folk music roots and influences. I just choose to show it for samba music.

The carnival in Rio has become a major commercial tourist attraction. But still it is a cordial matter to the *cariocas* (inhabitants of Rio). There are about fifteen samba schools in Rio. Each school prepares the appearance at the carnival for half a year. Thousands of members are engaged in song composing, costume making, and music and dancing rehearsals. The schools are located in and associated with the *morros* – the hills where the *favelas* (shantytowns) are built. These favelas have tribal features. They are illegal enclaves in the city with their own drug related gangster governments and their own gangster laws. There are violent feuds between the favelas. It is estimated that one third of the population in Rio (11 million people) live in such enclaves.

The samba school is important for the identity of the people at the morro. Every school has its own name, its own honourable history, its own colours, its own heroes, and its own rhythmic accent. This accent is very important. “It is death penalty to confuse those rhythms”, the samba music arranger Jota Moraes jokingly (?) told me (Moraes 1999)[317]. To play the wrong rhythm at the carnival in Rio could be compared with sitting in the *Vasco* section with a *Flamenco* t-shirt at Maracaná Football Stadium. It is not considered healthy. In fact the samba schools compete against each other just like football teams do. There is a league and the schools can move up and down in a system of divisions, depending on their performance at the carnival. While this official carnival is displayed for pay-

[317] Moraes, J. (1999) Personal information 21/5

ing audiences, there are also several unofficial carnivals in the streets where anybody can participate.

Thus for one week each year local self governed tribes wearing their own emblems, masks and dresses, interact totally entrained in a ritual where they play their own musical accent, dance their own movement pattern, and compete with other groups.

The connection between rhythm and tribal identity is not unique for the samba schools. We can see it in Afro Brazilian religions. In the four-day non stop *Congado Festival* in Belo Horizonte, Brazil, ten to twenty *Congado* groups, each with a rhythmic pattern of its own, a tempo of its own and a choreography of its own, confront each other (Lucas 2005)[318]. At this confrontation the drummers insist on keeping their own rhythmic pattern. They have to be synchronized with their own group and avoid synchronization with the other groups. This is difficult to do since we have a tendency for entrainment.

> These are the moments we need more strength and firmness, we have to be very careful and only look at the skin of our drum. (Geraldo da Luz, translated by Glaura Lucas 2002: 73)[319]

What then is staged at the carnival? In his analysis of the medieval carnival the Russian linguist Mikhail Bakhtin (1895–1975) recognized the *grotesque* (Bakhtin 1984). This word, Bakhtin explains, emanates from the Italian word *grotta*, alluding to rock-paintings in caves found in Italy. These are characterized by a free and creative play with plant, animal, and human forms transforming into each other – creating each other. The grotesque could be interpreted as a cosmic vision for life conditions; life is created from life, birth from death, and form from form. Bakhtin saw the carnival as such a grotesque unity, a dancing and twining body consisting of bodies – copulating, terrifying and hilarious.

---

[318] Lucas, G. (2005). An ethnographic perspective of musical entrainment. *In Time with the Music: the concept of entrainment and its significance for ethnomusicology* (Eds. M. Clayton, R. Sager, G. Lucas & U. Will. *European Meetings in Ethnomusicology (special ESEM-Counterpoint volume).* Vol. 11. 96–99

[319] Lucas, G. (2002) *Os Sons do Rosario: O Congado Mineiro dos Arturos e Jatobá.* Editora UFMG, Belo Horizonte

The grotesque captures separate moments of time in one image, presenting something that conquers logic. Time as a symbolic representation is dissolved into the present moment. The confrontation of two *cosmologies,* the cultural ideology of prevailing power and the grotesque, releases *a popular laughter*. It is the sudden insight that we are all the same. This shared "something" is nothing less than the universal conditions of birth and death, generations of decaying and renewal; in one word time – presented in the grotesque picture.

Carnival is, in the analysis of Bakhtin, the occasion when the people come out in the streets to play with the structure of power. In this sense the carnival is a ritual. Something is replaced, staged and played. It is a degradation – to bring down everything lofty, idealistic, and abstract to the level of the body. The official moralist discourse is inverted. Or even, as Bakhtin says, the central thoughts, images and symbols of official culture are transmitted to the lower bodily parts. In feudal European societies the official ideology was the Christian vertical hierarchical conception of the world with heaven and hell. Hell was visualized in masks and dramatizations, but, and this is emphasized by Bakhtin, in a playful popular way. Laughter conquers horror – demystifies horror. Laughter, Bakhtin claims, liberates us, "not just from the outer censorship but from the immense inner censorship". Bakhtin recognizes a renewal in this: death giving birth, affirmation of earth and body – of sexuality.

The mask has a connotation of change and reshaping identity. It is ridiculing in a play grounded in a very special connection between reality and image. In the menagerie of characters the *Harlequin* had a special role. With his padded cloths he personified the always hungry servant, cunning and witty. The word Harlequin can be deduced from *Erlekin* (devil, as in *Erlkönig* by Goethe). In the diableries in the medieval mystery plays, precursors to the carnivals, the devil was presented as the gay spokesman for the up-side-down holiness.

There is thus a transmission from cognitive consciousness to procedural (bodily) consciousness. The carnival stages, Bakhtin claims, a "carnivalization of consciousness". This is how Bakhtin depicts the "carnivalesque" crowd:

> This festive organization of the crowd must be first of all concrete and sensual. Even the pressing throng, the physical contact of bodies, acquires a certain meaning. The individual feels that he is an indissoluble part of the collective, a member of the people's body. In this whole the individual body ceases to a certain extent to be itself; it is possible so to say to exchange bodies, to be renewed (through change of costume and mask). At the same time the people become aware of their sensual, material bodily unity and community. (Bakhtin 1984: 255)

The carnival thus produces a shift from a cultural perspective to a tribal perspective.

The Brazilian anthropologist Roberto DaMatta refers to Bakhtin in his analysis of the carnival in Rio (DaMatta 1991). DaMatta recognizes the structure of power in Brazil – the *Casa*-principle. *Casa,* literally meaning house, refers to a feudal structure studied by the world famous Brazilian anthropologist Gilberto Freyre in the thirties (Freyre 1964)[320]. The patriarch at the *casa grande* (farmhouse) controlled his family, the slaves and the neighbourhood. This can be spotted in the history of politics in Brazil where populist dictators have taken on fatherly roles. Against this corrupted vertical structure of dependencies and bribes, DaMatta puts the horizontal *Rua (street)* principle. This refers to democracy – the anonymity of the street where everybody is equal – an anonymous nobody.

According to DaMatta the feudal *casa*-principle still governs Brazil. It constitutes a hindrance to democratic development. This, DaMatta states, is the Brazilian dilemma. Brazil is formally a democratic constitution – but in practice it is ruled by hierarchical connections, bribes and guns. Brazil is one of the most unequal countries in the world. At carnival the *casa*-structure is highlighted in a distorting mirror. Carnival is *loucura* (madness). The rich dress down and the poor dress up.

> During festival it was a fiesta and everybody was happy. You were not a washing woman or a worker or a bum. You dressed up as

[320] Freyre, G. (1964) *The Masters and the Slaves.* Borzoi Books, New York

> dignified people – counts and lords – as better people. Then court moved to the ghetto. (Old woman interviewed by Olsson 1977)[321]

But in the Carnival people also expose themselves naked. They show off sexually. Samba is the dance of the hips.

> Carnival is great, to let out the bicho (wild animal), to be loose. It is the time of fantasia (fantasy), as they say. It is called fantasia to dress up. But in fact it is not. In fact we are masked the whole year but not during festival. (Gonzaguinha 1979)[322]

In this quote Gonzaguinha (mentioned in the introduction as an inspiration to this work) expresses Bakhtin's idea. All the year there is a cultural repression. Only during carnival the people are free to express themselves. The tribal mechanisms produce this freedom. This contrast is particularly accentuated during times of hardship.

> Could you imagine a country with the problems we have got if there was no alegria (carnevalistic happiness), joy and music? It would have exploded. People would have killed each other long ago. Without music this country wouldn't have existed. (João Bosco in Bar Brasil 1990)[323]

The unification of individuals in the crowd is reflected in the music. A carnival samba orchestra consists of hundreds of individuals like a great symphony orchestra with the difference that everybody is playing percussion. It is a web of rhythm sections creating a breathtaking polyrhythmic mass effect. Some of the sounds are clearly carnevalesque as the agogô sounding like a cow-bell, the cuíca performing crazy sounds, and smaller bells with their connotations of jesters. Rhythmic individualities and personalities become scales in this colourful anaconda.

---

[321] Olsson, L. (1977) *Samba*, TV documentary produced by Lennart Olsson for TV 1 Sweden. Broadcasted March 2

[322] Gonzaguinha (1979) interview by Radio Globo.

[323] Bosco, J. (1990) Interview in Bar Brasil, a Swedish Radio series on Brazilian music by Lennart Peck, Lennart Kjörling and Sven Kristersson, part one. Broadcasted July 2.

The elements that we recognize in the Carnival in Rio de Janeiro can be spotted in carnivals all over the world – with variations depending on the culture structure the carnival is confronting.

In Brazil as in the rest of the world samba is associated with the body and the people – yes to the body and yes to the people. In the Samba, *É* Gonzaguinha stages a crowd demanding recognition: "Look at us, we do not mind sweating, but from pleasure" (Gonzaguinha 1994)[324]. The message is: We, the African slave descendants, are not just a working force but humans. And how appalling this may seem to you, we are not just manpower, we are entitled to sensual pleasures.

Sangrando

The ritual is a forestage to art (Dissanayake 1992). In art we can see, as did Lévi-Strauss, the manifestation of myth. This is not only the case in visual art but also in poetry. A beautiful example is Gonzaguinha's song *Sangrando* (Bleeding)(Gonzaguinha 1989)[325]. In this song Gonzaguinha expresses how the relation between him and his audience in the very act of the performance becomes charged with mythological meaning. He is conscious about this meaning and communicates this consciousness to the audience. This is, I would say, an important aspect of the "commission" of the artist – to imagine the mythological meaning of the situation, and to catch this imagination in a form (in the case of poetry a word-form). It is to some extent to move tribal consciousness to the realm of cultural consciousness.

Gonzaguinha (1945–1991) was raised in favela São Carlos in Rio de Janeiro by foster parents. He identified himself as a *moleque* (a pejorative name for street kids). In the eyes of the poor, he was a hero because he, the successful artist, was originally one of them (African heritage, poverty, and the bodily felt experience of struggle for life day by day in the streets). A contributing factor to his high esteem among the people was his courage. He openly criticized the

---

[324] Gonzaguinha (1994) CD: É. *Mestres da MPB*, Warner

[325] Gonzaguinha (1980) CD: *De Volta au Comeco,* EMI

feared and hated military dictatorship and was repeatedly thrown into custody.

I never had the privilege of hearing Gonzaguinha live, but I witnessed a concert in Rio de Janeiro 2004 where the singer/actor *Gaspar Filho* impersonated Gonzaguinha and performed his songs, among them Sangrando:

Bleeding

When I release my voice
Please understand
That every word
Is someone opening up
The heart in the mouth
Bleeding
It is the battles of our life
I am singing

When I open my mouth
This strong force
All you can hear
Is surely lived
See the lustre in my eyes
And the trembling in my hands
And the sweat of my body
Carrying all my race and emotion

And if I cry and the salt damps
My smile
Don't be surprised, the song
You are singing
Is my strength to sing
When I release my voice
Please understand
This is just my way to live
And to love

(Interpretation to English by the author)

The thin body is crying out: Look at me, I am singing my life, my anger, my sufferings and my love. I give you my heart. I am bleeding. Almost every line contains body metaphors. What happens to metaphors as we entrain the music? Let us once again bring Merleau-Ponty's words about the myth to the fore:

> Myth beholds the essence within the appearance, the mythical phenomenon is not a representation, but a genuine presence. The daemon of rain is present in each drop…the soul present in each part of the body. Every apparition is in this case an incarnation (Merleau-Ponty 1996: 290).

In the tribal perspective the metaphors appear as a "genuine presence" – thus as a perceived phenomenon of lived life. As we become caught in the present moment, symbolization ceases to exist. The cognitive consciousness is replaced by phenomenal consciousness. The words will be bodily experienced. The words become flesh. The myth is incarnated.

Interestingly the Brazilian word for song "*canto"* also means place or corner. So the line *your song is my power to sing* could also be translated: Your standpoint is my power to sing. Given this connotation Gonzaguinha expresses that he takes the perspective of the audience and this gives him the power to sing. The whole stadium is electrified. The power from the audience is fuel to the artist. Gonzaguinha defines his relation to the audience when he connects his bodily experiences with the experiences of the audience: "It is the battle of *our* lives I am singing".

We follow the music from the soft melancholic first statements, where Gonzaguinha in a spoken tone merely reflects on what happens when he sings, to an intensification one octave up when the speaking converts to singing and the text suddenly reflects the present situation: "See the trembling in my hands". As the song unfolds, the presence introduces itself in the lyrics.

The emotional climax comes on the word "race". The Brazilian word for race *(raça)* is hard to translate. It is metaphorical with connotations ranging from instinctive commands of the body to the call from the African blood. It is referring to the power of the singer. When Gaspar sings *All my race and emotion* the Brazilian pronunciation of *raça* with the hissing "r" from the depth of the throat hits us like a snakebite. The presence becomes acute. We feel the raça in our own skin. Now his throat is our throat, his chest is our chest, his sweat is our sweat and his heart is our heart.

I experienced this. The reactions of the audience affected me. There was this force touching us. Some were crying. All of us were deeply moved. At this moment the singer was connecting the hearts in the audience.

There is an element of African rituals in most Brazilian religions. In these the proselytes, dancing to the sound of drum rhythms, are possessed by spirits. There is thus a disposition in the audience to accept such transitions.

This is not to say that the audience is possessed by the spirit of Gonzaguinha, when these words hit them in the music, neither that this street-smart and sometimes aggressive singer from the favela was some kind of saint. Far from it. Rather that the performance is saturated with the myth. Gaspar sings Sangrando holding his hands out as to meet and embrace the audience. Not occasionally, but through the entire song. This is an iconic sign the *cariocas* (the inhabitants of Rio) can recognize – *Corcovado.* This gigantic monument of Christ placed on a 600 m hill can be seen from any angle in Rio as if blessing this criminal city in the tropics. It is blessing and crucifixion in the same image. Bleeding wounds are opened, not necessarily consciously associated with Jesus. But the image charges the words as batteries are charged with electricity. When Gaspar sings about the sweat transporting all his race and emotion, we are on line as if wired to one another and the shockwave from these charged words hit us physically. At this moment the myth Gonzaguinha is mediated by Gaspar. We live the myth. We *are* the myth.

The artist is personifying our suffering on stage, handing over his body to us. It is revolutionary in the sense that Jesus Christ was revolutionary, when he suffered for us and at the last supper offered his body and blood in the communion. "*This is my way to love*". The essence of the communication is this flesh and blood. This is us. This is what we are: This unbearable division in divinity and fragility - this unbearable longing.

The rock concert

There is expectation in the air. We can feel the sensibility of the audience. Sometimes this can be a concrete experience, as can be understood by formulations as: "the atmosphere was so tense that you could almost touch it". When the band enters the stage the expectation is released in a unison tribal roar.

We forget about ourselves. The united crowd moves with the artist. It is one single body. The listeners merge with the band. And with each other. This can be seen in the perhaps somewhat outdated practice to use a lighter at emotional moments. As if to show that their hearts are glowing, people wave their lighters. A wave of thousands lights shows that we are all connected to this e-motion.

In an interview with Jerry Garcia and Phil Lesh in *The Grateful Dead* David Gans (1991)[326] tried to catch what happens between the members in the band and between the band and the audience at a concert. Gans start with a question concerning the mythology of the band.

> **Gans:** How did the personal, collective quest turn into this incredible myth? […] When did it happen? How did it get transferred to the crowd?
>
> **Garcia:** I don‘t know. It seems to me that as soon as we were playing at the Acid Tests, you‘d look out there and you‘d see that guy and he‘d look up and go, "Yeah, I know what you guys are doing. I know what you guys are up to!" And you knew that they knew. It was one of those things. It was like one on one. Recognition, it was flashes of recognition. No telling what, really, but that thing that you recognize. And you recognize it because it‘s there. I could never say what it was.
>
> **Lesh:** It‘s communal, in the sense that it exists—
>
> **Garcia:** In that moment, when that guy looks at you and goes "Yeah, I get it".

[326] Gans, D. (1991) Conversations with the dead. *The Grateful Dead Interview Book*. Citadel Underground, New York

Lesh: You‘re both in the same place.

Garcia: That‘s right. You know, and he knows, and that‘s that moment.

Lesh: It‘s as close as you can come to being somebody else.

It is interesting to follow the discussion about what this “it” may be. Obviously it is a real experience, but it is very hard to nail down what “it” is. It is not a semantic experience. What is agreed on here? It is communal, they say. It is some insight that is recognized. It is confirmed in the eyes of someone in the audience.

Then Lesh says: *You’re both in the same place*. This is the key. It is the shared perspective. And it is so astonishing that these are his exact words, because everything else is vague. The shared perspective of what? This question is difficult to answer because the music does not signify any shared meaning. But remember that the whole conversation started with the search for the mythology of the group. This is “it”. This is what is heard in the musical interaction – I am part of it. We all are.

It is unthinkable that a single spectator would act the same way a spectator in the crowd does in the rock concert. The base drum is insisting: NOW, NOW, NOW until everybody is linked to the same now. A drumbeat is not a monotone copy of something already heard and known, but every time a unique event enriched by all the other drumbeats. Where the analytical listener feels bored of the monotonous repetition, the listener entrained is captured in the present moment. There are in his mind no former beats. There is just this one beat connecting him with thousands and throwing his head more violently at each individual beat as if this particular beat was the ultimate experience. And this very beat explodes as he leaves the earth with a jump in a sustained n-o-o-o-o-w. And there he sails in a body no longer named Patrick, or Betty, or Bob – but *Weeee*. And the drum is no longer on stage but beats inside *Weeee*.

And when we leave the concert, we can see in the eyes of totally unknown people that we have something in common. We have shared the experience, although we do not know, as Lesh said later in the interview, “whatever the fuck it is”.

Listening alone

Can a CD bring forth a tribal perspective in a sole listener? As already pointed out communal speech and singing engages the planum temporale. This area is not activated when we speak or sing alone (Saito *et al.* 2006). Since this area is also activated when we just listen to music, I draw the conclusion that music listening entails synchronization with the musicians. As we start moving, we take on a jazz, a punk, a flamenco etc. attitude to the world. This is, as if we were to spell this out, how we dance the world – and make it appear. Listening to music we are always embraced by the tribe. Listening to music, we are never alone.

Discussion

It is a wonderful experience in music to share the rhythm of time and movement with others. This is to intuitively feel where the band is heading; to become part of one whole music body and to know how to contribute.

Different instruments take on different roles. I would even say different *social* roles. The drums lay out the common time perspective, the base provides the solid structural ground, the keyboard colours the space with harmonies and the soloist has a free role of emotional expression within the laid out frame. These instruments attract different personalities. Musicians are group people; bassists more than others. They live for the group, whereas soloists are more individualistic. They have to be. The soloist must love taking up space, attract attention, exhibit emotions, while the others must love backing her; even to the limit of provoking her, exciting her and moving her – literally. In music as well as in any tribe or society the perfect mix is needed. Or shaped by the group. Interacting groups shape and distribute roles. And individuals have a tendency to live up to the roles attributed by the group.

I suggest that this is connected with the ritualistic element in music, since we know that rituals define and redefine roles.

As mentioned in the introduction of this chapter David Huron has suggested that the evolutionary explanation of the music phenomenon lies in its capacity to create bonds. I agree, but I think that this becomes even more obvious if we consider the ritual rather than the music as the original phenomenon. Most anthropologists agree that music as a universal phenomenon is ritualistic. It is only in the modern era music has developed as a separate art form. Freeman (2000)[327] discusses this in his article *The neurological role of music in social bonding*. Dance is keeping together in time, he claims. "Ritual music and dance foster social bonding and so has been essential to creating the trust upon which all social interaction depends".

Music is a part of a multimodal phenomenon. And it is perceived as such. Mirror neurons are multimodal (Keysers *et al.* 2003) and music, contrary to most other activities, engages the whole brain. It is not just the music we copy, it is the whole act. And we are participants in that act.

> Simply, music moves us. When music engages the human mind most strongly – when performers play music, or when listeners tap, dance or sing along with music – the sensory experience of musical patterns in intimately coupled with action. Thus, music is an excellent example of a perception action cycle… (Janata & Grafton 2003)

At a concert we are not just interested in tone relations but in what the artist is expressing on stage. Even at home, as we listen to a record, it is not only the sound that is of interest. We can be totally engaged – we move, we feel, and we imagine. We have a relation to the group and the music style and our identity is coloured by that relation.

It is tempting to agree with Huron about the evolutionary benefits of music to the tribe, but it is even more obvious that the ritual is beneficial, since the ritual is communicating knowledge; it is communicating a way to perceive the world. Through the rite humans

---

[327] Freeman, W.J. (2000) A neurological role of music in social bonding. *The Origins of Music* (Eds. N.L. Wallin, B. Merker & S. Brown). MIT Press, Camebridge, MA. 411–424

could share experiences. Evidently those tribes who developed this skill had an advantage.

Many scientists have speculated about the development of music. Let me just present a thought on how things just might have developed:

1. A returning hunter communicates his experiences by way of a concrete showing, mimicking the actual moves and sounds of the hunt.
2. Spectators chose to interact in this staged hunt and follow in the steps of the hunter. This imitation capacity is inborn. The participants recognize in their own bodies how the hunt took place.
3. The ritual is performed *before* the hunt, just to show the tribe that today is hunting day.
4. The ritual is performed to make the hunt successful, because last time it became successful after this ritual. The ritual now has tribal meaning. It means hunting, it means we are the hunters, it means today is hunting day and it means: If we do not do this, the hunt will not succeed.
5. Rhythmic performances make the ritual easier to entrain.
6. Scales make the sounds easier to entrain.
7. Those tribes who developed rhythm and scales survive because they can communicate better than other groups.
8. The music is separated from the ritual to be developed into an exquisite art form in its own right, playing with the tools developed in rituals.

If the ritual is the origin of music, it explains why music is never perceived just as music but as movements, visions, emotions etc. We perceive the rite in the sound of music.

As mentioned in the introduction, a classical question in music philosophy is: Why do we listen to sad music? Why do we at all engage in other people's problems? There is, from an egoistic point of view, not much reward to be had from this. There is possibly a gain in a group perspective, that we have the ability to be empathic. This depends on the individual's ability to perceive things from a wider perspective. As we have seen there is an altruistic streak in groups. We should expect that some kind of emotional reward is implicated.

Funerals are perhaps the oldest and most universal of rituals. It is easier to understand why a funeral releases sadness than why music would. Sadness is an emotion associated with loss of relation – with longing. Sad music mimics mourning. It makes us take part in communal mourning; interact in mourning. We are linked to a superconductive emotion. Music is the outstanding medium to express sadness and longing in love songs, religious music, and all kinds of nostalgic songs about home, mummy, good old values, etc. It is the most arousing emotion in music (Panksepp 1998). This is not only reported by Panksepp but also in Bossius's interviews on trancing (Bossius 2003)[328] and in Becker's investigations of deep listening Becker (2004).

Gabrielsson & Lindström report that relational emotions are commonly reported at strong music experiences, particularly emotions of safety and warmth: "I felt safety and community with all who participated" (Gabrielsson & Lindström Wik 2003). This emotion has a counterpoint in neurophysiology (Blood & Zatorre 2001). It is established that pleasurable responses to music correlate with activities in the brain regions implicated in reward. Opioids are released by social contacts, sex and play – and by music.

It is just possible that ritualistic group mourning produces an emotional shield. As we take part in a funeral, as we collectively cry out the loss, we are comforted by the interaction of the group. The interaction in itself produces warmth and safety. It creates bonds that make it possible to endure as a group and move on together.

Does then sad music produce opioids? Sadness occasionally causes "chill responses" sometimes called *piloerection* (Panksepp 1995)[329]. A chill is felt, the hair on the arms is erected, a sensation is felt down

---

[328] Bossius, T. (2003) *The Future in the Rear Mirror. Black metal and transculture. Young people, music and religion in a post-modern world.* Daidalos, Gothenburg

[329] Panksepp, J. (1995) The emotional sources of chills induced by music. *Music Perception*. Vol. 13. No 2. 171–207

the spine and tears may come into our eyes. It is a warm experience. Such chills are thought to be caused by opioids.

Even if the listener is alone and listens to sad music, she takes part in something that is perceived and felt as a ritual. This connects her to her bodily experiences of loss and to episodic memories of loss. But in this the warmth from invisible others is felt. This fills existential emptiness with a meaning that is not, and cannot be, formulated in words.

Maybe the question why we are emotionally affected by music should be rephrased: Why are we emotionally affected by rituals? This question is easier to answer. Rituals are not abstract tone relations. They are human interactions.

6

Contemplating Music

The Cultural Perspective

On its way to the listener music passes many transformations such as the music culture that surrounds and inspires the composer, the fantasy and the thinking of the composer, the encoding to a music score, the decoding by the musician of the written music, the musician's interpretation of this decoded music, the sound of the instruments, the phonogram, and the perception of the listener. This could be divided further. On its way from the composer to the listener the music occurs in different shapes. We can see all these stages as signs of each other – as representations. Where then is the music? Or more precisely: What is a sign of music and what is music? We could also ask in what sense music could be a sign of extra-musical reality. I intend to show that the answers to these two questions are perspective dependent.

Depiction of perspective

Michael Gazzaniga describes in *The Social Brain* (1985) how his split-brain experiments led him to the conclusion that the human brain has a module in the left hemisphere with a special and specifically human capacity. He called this module *the interpreter.* This device allows us to interpret our perceptions and to interpret our actions (rationalizations). Cave paintings, graves and mythology inscribed in tools bare witness of early human efforts to interpret her existence. From there on perceptions could be taken as signs of something. The most important example of this is the coded sound – the word.

With the interpreter it is possible to communicate without bodily interaction. The *code* (a conventional connection between a sign and its denotation) is a prerequisite for accumulation of knowledge. It makes mass communication possible. And it allows for knowledge to be passed on to the next generation. The code, in other words, shapes larger entities of people with common understandings – *cultures.*

The concept *mentifacts* designating intellectual, aesthetical, spiritual and ethical values in a culture was introduced by the English biologist Sir Julian Huxley, along with the concepts *sociofacts* for social structures and *artefacts* for material objects (Huxley 1951)[330]. Roland Posner (1989: 265)[331] treats mentifacts as a *set of codes*. Posner proceeds: "Mental culture can therefore be regarded as a set of codes". This leads to a semiotic definition of a cultural society as a set of code users. Anybody in possession of the cultural codes belongs to the society. I will discuss "culture" in this narrow sense.

If the codes are shared in the culture, we have a language. This language does not have to be verbal. People create in their daily communication semantic fields. The discourse produces the meaning of the concepts of the culture. When we use codes, we see the world through the eyes of the culture. I will call this *the cultural perspective*. This delimits my definition of the cultural perspective to the cognitively conscious.

In the cultural perspective we take an indirect attitude to the world. We deal with the representations of the world, not with the world itself. We use signs. The sign is the vehicle to leave the here and now. With the sign we can think about things in the past and things to come. We are no longer locked to phenomenal direct experience of the world. We can think about things that are not here

---

330 Huxley, J. (1951) Knowledge, morality & destiny. *Psychiatry*. May 14. 129–151

331 Posner, R. (1989) What is culture? Toward a semiotic explanation of anthropological concepts. *The Nature of Culture* (Ed. W.A. Koch) Brockmeyer, Bochum. 240–295

and things we have not experienced but just read about. We can share other people's thoughts, the thoughts of an author and the commonly shared beliefs of the culture. With the sign "me" I can step out of myself and think about myself as an agent. I insert the current discourse of individuality. For example I can think about myself as someone who listens to music. This is totally different from the co-perception of myself in music listening.

Music and the cultural perspective

Music is contemplated, reflected on, thought about, analysed and talked about by musicians, listeners, critics, and scientists. Music is inscribed in culture, in a weave of texts – a discourse. It reflects culture and is thus – in itself – a sign of culture. Music is artefacts as products, sociofacts as institutions and mentifacts as discourse.

Musicology, as all sciences, always takes a cultural perspective, simply because it uses cultural codes and mentifacts. I want to emphasize that this is a perspective because you often see that musicologists generalize the scientific analysis of music to the perception of the music listener. This is a mistake. The perspective of the listener is, as we have seen, not always cultural. Musicologists traditionally work with the analysis of written music. This is far from dancing, everyday listening and tribal experiences. Since the bodily interaction with music is not part of traditional musicological method, we tend to underestimate this side of music perception.

Cultural knowledge about the composer, the artist, and the music, is undoubtedly important for everyday listeners. All kinds of aided guidance, such as programs, song lyrics, interviews, and lectures direct the understanding of a piece. No doubt biographical knowledge about the composer or the artist contributes to the emotional experience as we listen to music. The information that Eva Cassidy died of cancer shortly after having recorded *Over the Rainbow* affects my understanding of the lyrics, the voice and the phrasing.

But *how* does it affect emotion? We cannot exclude the possibility that our explanations why we are moved are not always true. As we have seen from the neurological evidence, the emotional reaction sometimes comes first. Sometimes the context is not the cause of the emotional reaction. It helps us to explain the reaction in words, but this explanation is in fact a rationalization. "Why do you like this music?" we may ask. And the respondent might answer that the voicing is excellent, that the singer is charismatic, that the sound is beautiful, that this whole genre is interesting and so forth. Why is this judged good? *Because we like it.*

What, then, comes first, rational judgement or emotion? This is a crucial question. We will come back to it in a chapter solely devoted to the discussion of emotion.

This discussion leads to the following questions:

1. Does music signify contextual meaning?
2. Does this signification produce emotions in the listener?
3. And ultimately: Can music signify emotions?

These questions cannot be answered without a review of semiotics and what this science actually claims and can claim.

Semiotics

The ancient Greeks thought about symptoms as signs of illness. Since then representation has been a subject of interest for philosophers such as Augustine (354–430), Thomas Hobbes (1588–1679), and John Locke (1632–1704). Modern semiotics was fathered by the Suisse linguist Ferdinand de Saussure (1857–1913) and the American philosopher Charles Sanders Peirce (1839–1914). They were followed by Ernst Cassírer (1875–1945), Louis Hjelmslev (1899–1965), Charles W. Morris (1901–1978), Roman Jakobson (1896–1982), and Umberto Eco (b. 1932), among others. The aspects of this field are covered in an anthology consisting of four enormous volumes called

Semiotik – ein Handbuch zu den zeichentheoretischen Grundlage von Natur und Kultur (Posner *et al.* 2004)[332].

Non-linguistic use of semiotics was introduced by structuralists as Claude Lévi-Strauss (b. 1908) and Roland Barthes (1915–1980), as well as by psychoanalysists as Jacques Lacan (1901–1981) and Julia Kristeva (b. 1941). Semiotics became the instrument to "read" a culture through its artefacts in anthropology to Clifford Geertz (1926–2006). One such artefact is of course music. For a theoretical understanding of the semiotic analysis of culture, see *The Nature of Culture* (Koch 1989)[333].

In the fifties semiotics made its entry to ethnomusicology and music anthropology as a cross-cultural method to analyse "utterances" of music from different cultures (Monelle 1992: 28). It developed into a tool to analyse the "meaning" of music (including affective meaning) in the 70s. Music ultimately was seen as a language with an explicit semantic meaning, which had to be decoded.

There is a vast literature on signification in music, e.g. (Nattiez 1975)[334], (Stefani 1976)[335], (Tagg 1979), (Martinez 1991)[336], (Tarasti 1994)[337] and (Cumming 2000)[338]. The principal lines are summed up by Monelle (1992).

There is not one semiotic theory, but many. Peirce's semiotics is the most elaborated system of them all. This is the source of commonly used concepts as the Symbol, the Icon and the Index. It is a general theory of all forms of representations. It has become widely applied in other sciences. Since this theory is the most influential in musicology, I will focus on that. Even so, there is the problem that

---

[332] Posner, R., Robering, K. & Sebeok, T.A. (2004) *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlage von Natur und Kultur.* De Gruyter, Berlin; New York.

[333] Koch, W.A. (1989) *The Nature of Culture.* Brockmeyer, Bochum.

[334] Nattiez, J-J. (1975) *Fondements dune Semiologie de la Musique.* Union Générale d'Editions, Paris

[335] Stefani, G. (1976). *Introduzione alla Semiotica della Musica.* Sellerio, Palermo

[336] Martinez, J.L. (1991) *Música & Semiotica: Um estudo sobre a questão da presentação na liguagem musical.* Potifícia Universidade Católica de São Paulo, São Paulo

[337] Tarasti, E. (1994) *A Theory of Musical Semiotics.* Indiana University Press, Indianapolis; Bloomington

[338] Cumming, N. (2000) *The Sonic Self – musical subjectivity and signification.* Indiana University Press, Indianapolis; Bloomington

Peirce, although the theory as such was fairly consistent over the years, never stopped redefining key concepts. For example it has been estimated that the key concept "sign" has 76 definitions. The many definitions are due to Peirce creative personality. In his many articles and letters Peirce always seemed to extract new aspects. This is enriching to the reader. It specifies Peirce's thinking, but it calls for care and knowledge about the whole work for followers and users.

It would have been nice to delimit the discussion to the question: Does Peirce's classification of signs in 1903 (when he published his work on signification) allow for music to signify (Peirce 1967)[339]? Even so, the question is difficult to answer. The reason is: I cannot claim to understand the theory. Even if I did, it would be an enormous undertaking to explain it. And even if I could, nobody, or almost nobody, would understand the explanation. In fact it is somewhat of an intellectual sport to depict practical applications of Peirce's sign classes.

We are in good company. Many renowned semioticians have got lost in their attempts to follow Peirce's writings. According to Pape (2004)[340], no one, "with the notable exception of Savan", has understood the inferential connections between sign classes and the categories (leaving at least one more successful interpreter). But what if all the others were right and Savan and Pape were wrong. Maybe it cannot be understood. There is a possibility, and this is a blasphemous thought, that the reason why the theory cannot be understood is that the premises are wrong. It is a problem that Peirce's philosophy, although visionary and way ahead of its time, is about one hundred years old. Accordingly there is new evidence in relevant fields not accounted for by the theory.

Peirce believed that semiotics was the answer to how we acquire knowledge of the world. Semiotics was just the means in a much greater undertaking. The ambition was nothing less than creating an

339 Peirce, C.S. (1967) *Annotated Catalogue and the Papers of Charles Sanders Peirce* (Ed. R. S. Robin). Amhearst, MA

340 Pape, H. (2004) Peirce und seine nachfolger. *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlage von Natur und Kultur* (Eds. R. Posner, K. Robering & T.A. Sebeok). Teilband 2. De Gruyter, Berlin; New York. 2016–2040

understanding of how knowledge of the world is possible. The ontological question was: How do we make up a representation in our minds of what is out there? Peirce tried to grasp the general structure of experience (Pape 1997). The aim was a *grammatica speculativa* that could describe the forms of all types of representation and knowledge.

One keystone in Peirce's philosophy was phenomenology, but this is the phenomenology before Merleau-Ponty. Phenomenology was in the air (James 1884)[341]; (Bergson 1889)[342] and Husserl (in extenso). Peirce criticized, but was never the less part of Cartesian Philosophy (Ochs 2004)[343]. The barrier between man and world creates the problem how we get access to this world. This was the problem Peirce tried to solve. This may be one of the premises that could be questioned.

With *The Phenomenology of Perception* (1945) Merleau-Ponty presented a new understanding of perception. The objectified body was subjectified. We interact directly with the world. And perception emanates from these actions. The world-mind duality was dissolved. Take a look at this statement by Peirce:

> It is that the word or sign which man uses is the man himself. For, as the fact that every thought is a sign, taken in conjunction with the fact that life is a chain of thought, proves that man is a sign; so, that every thought is an external sign, proves that man is an external sign. That is to say, the man and the external sign are identical, in the same sense in which the words homo and man are identical. Thus my language is the sum total of myself; for the man is the thought. (Peirce 1931: Vol. 58 § 5)

Signification is thus, according to Peirce, the only source of information. All mental activity is made up of signs. And consequently humans are signs. It could be noted that Peirce does not even bother

---

341 James, W. (1884) On some omissions of introspective psychology. *Mind*. Vol. 9. No 33. January. 1–26

342 Bergson, H. (1910) *Time and Free Will: An essay on the immediate data on consciousness*. George Allan and Unwin, London

343 Ochs, P. (2004) *Peirce, Pragmatism and the Logic of Scripture*. Cambridge University Press, Cambridge

to specify this sign to the human *thoughts,* or the human *mind,* but to *humans* as if there is no such thing as a body.

This is not commensurable with perception-action theory. Here are some of the clashes emanating from this false premise:

1. The separation of the world out there from our inner world must be rejected. The critique of the dyadic Cartesian ontology is an acknowledged consequence of Merleau-Ponty (1996) and James Gibson's (1986) phenomenology of perception. This is furthermore the message in Lakoff & Johnson's *Philosophy in the flesh* (Lakoff & Johnson 1999)[344]. The world, as we know it (and what else is there to discuss?), is, as we have discussed in the preceding chapters showing itself to us as a the result of action and interaction. Everything we know about it can be traced back to bodily experiences.

2. We do not *interpret* information to the extent that Peirce believed. Most understanding is automated in perception. A child does not interpret the bouncing of a ball as a sign of gravity; she can catch it anyway. Infants do not interpret the mouth movements of the mother as the consequence of muscular activity; they can imitate it anyway. Animals do not interpret situations; they act on them anyway. Intentional meaning and situation understanding are thus integrated parts of perception. We do not decode stimuli but perceive the situational meaning directly as "affordances", which means that the use aspect is evident. We perceive emotional information from an utterance without decoding it. This information is in the tone of voice – the emotional prosody. In emotional face-to-face interactions we perceive each other and understand the social situation directly.

3. A perception is first of all an invitation to act. We are not likely to interpret something coming in our direction as a ball before we perceive it as something to catch or duck.

[344] Lakoff, G. & Johnson, M. (1999) *Philosophy in the Flesh. The embodied mind and its challenge to Western thought.* Basic Books, New York

The invitation to catch the ball is automated in the perception.

4. Thus we must distinct phenomenal consciousness from cognitive consciousness. Cognition is semiotic; perception-action is not. Phenomenal consciousness is typically situation dependent. Gazzaniga (1985) provided neurological proof with his series of split-brain experiments that we in fact are capable of acting without any foregoing rational decision. Rather we tend to invent reason after action. Phenomenologists therefore use the concept perception to imply a being-in-the-world perspective – the perspective of the present moment. Perception is prior to semiosis.

5. Knowledge does not have to be encoded in the semiotic sense. Knowledge of how to perceive learnt by perceptual learning, procedural knowledge, and episodic memories are three examples of knowledge that is not semantic.

6. Neural representations are not signs to the perceiving subject (but to the neurologist).

All these points falsify the thesis that humans can be reduced to signs. I am not just a sign. I am also the phenomenological body, co-perceived, interactive, and intentional. Peirce never stopped working on the semiotic project, indicating that even to the originator there were always loose ends. One such loose end is the concept *Firstness* described as a pre-reflective quality. Icons (see below) are said to have Firstness. The idea of a pre-reflective sign is an anomaly, as I intend to show.

Since the work was never completed and since key concepts have various definitions, followers tend to present their own view. They do so without reference to the systematic layout and disciplines in Peirce's philosophy (Eco 1979[345]; Schönrich 1990[346]).

---

[345] Eco, U. (1979) *The Role of the Reader*. Indiana University Press, Bloomington

[346] Schönrich, G. (1990) *Zeichenhandeln: Untersuchungen zum Begriff einer semiotichen Vernunft im Ausgang von Charles Sanders Peirce*. Suhrkamp, Frankfurt am Main.

To sum up:

1. Semiotics is not one, but a set of theories.
2. The most used theory is not internally consistent and has never been completed.
3. Peirce's theory is, although based on phenomenological theory, not commensurable with perception-action theory.
4. Followers do not understand the foundations.
5. Followers make up their own versions.

The sign is perspective dependent

The simplest understanding of the sign is the dyadic constellation signifier/signified for the representation (representamen) and the object signified. Peirce's most influential contribution is the discrimination of three classes of sign relation (the relation between the sign and the object signified):

1. Icon. The signifier has some characteristic in common with the object signified "(…) by partaking in the characters of the object". (Peirce 1931–1958 vol. 4 § 531). This could be resemblance (e.g. a painting of a falcon).
2. Index. This sign is caused by causality or closeness (Peirce 1931–1958 vol. 4 § 531) (e.g. a falcon feather).
3. Symbol. The Symbol is a conventional sign (Peirce 1931–1958 vol. 2 § 307) (e.g. the word "falcon").

By "conventional sign" we mean that it is made a convention, i.e. the meaning is agreed and generally accepted. To say that the Symbol is conventionalized thus means that it is culturally coded, as are words. The understanding of a Symbol thus implies a cultural perspective.

I will now argue that to detect a sign, any sign, we need to take a cultural perspective. I will use the 1897 definition of the sign. This definition is chosen because it is the most well known, because it approximately covers most other Peircian definitions of the sign and because it contains the *interpretant* which is an important agent in Peirce's triadic understanding of the sign:

> A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the interpretant of the first sign. The sign stands for something, its object. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the ground of the representamen. (Peirce 1931–38: vol. 2, § 228)

This definition makes all signs cultural. I have two arguments for this:

1. The sign "stands to somebody for something". This means that the beholder must be able to discriminate between the sign and the "something" that it represents. If the perceiver is not aware of the difference between these two entities, she will have a phenomenal perception. The "something" will be perceived directly. When we smell a hot dog, our salivation turns on; we feel hungry and we start looking for food. Any meat-eating animal would do that.

2. The *interpretant* – the "equivalent" and "more developed" sign "in the mind of that person" is created by the original sign. This is the thought of the sign. This interpretant must, as for reasons already stated, be able to separate the two constituents of the first sign. There must be a border between these two constituents. According to Louis Hjelmslev (1953) meaning gives substance to form. This is the form of content. There is, he states, no possible existence for this meaning unless it has a form. The "paradigm" in one language and the corresponding "paradigm" in another language cover the same zone of meaning, which without these languages is an undivided, amorphous continuum. The borders in this zone of meaning thus do not appear

until they are formed by language. This demands a cultural perspective.

Perhaps it could be objected that Hjelmslev must be wrong, since we can think in images and separate the image from the reality. However in Peirce's understanding of iconicity does not permit that:

> So in contemplating a painting, there is a moment when we lose the consciousness that it is not the thing, the distinction between the copy and the real disappears, and it is for a moment a pure dream – not any particular existence, and yet not general. At that moment we are contemplating an icon. (Peirce 1885: 181)[347]

This depiction of the icon does not fit the definition of the sign. If the border between the copy and the real disappears, the beholder cannot possibly perceive that something stands for something. This means that to the viewer there is no sign. The content is seamlessly merged into the form. As soon as I say to myself: This painting depicts a falcon I take a cultural perspective.

This is why I stated earlier that a pre-reflective sign is an anomaly. And this is why I argue that signs are perspective dependent. Let me exemplify:

1. Tommy is thirsty and sees a glass of water. He gets an immediate urge to drink, which is coloured with the emotional aspect of thirst and quenching thirst. The glass in this case is not a sign of something to drink. Rather the use aspect, as James Gibson pointed out, is immanent in the perception. The urge to drink is unmediated. Now Tommy's sister Mary makes her entrance. Mary thinks (verbally): "Tommy is looking at a glass of water. Maybe he will have a sip." For her the existence of the glass and the act of drinking is not the same thing. In the first case we have a perception and in the second case we have signification. In the first case the glass does not stand for something, while in the second case the glass does stand for something. Thus

[347] Peirce, C.S. (1885) On the algebra of logic. A contribution to the philosophy of notation. *American Journal of Mathematics*. Vol. 7. No 2. January.

the same glass of water can be a perception including the impulse to drink in one perspective and an interpretant consisting of a mental representation of the sign *glass/drinking* in another perspective. In Mary's perspective the glass and the possibility to drink are connected "in reference to some sort of idea" (as the definition reads). It is not until this idea is developed in Mary's mind that the glass/drinking sign appears.

2. Exposed teeth in wolves may be interpreted as an indexical sign of aggression. By humans. Not by other wolves. To them it *is* aggression. To talk of exposed teeth as semiotic communication is to take an anthropocentric perspective. Only humans can be stupid enough to reflect on the meaning of the exposed teeth when confronted by a wolf in rage.

3. Let us say that we *smell* a hamburger. Most of us would agree that this is an indexical sign of a hamburger. But what if we *look* at it and draw the conclusion that this is a hamburger. This is the same thing although in another modality. Still we would say that this is a hamburger rather than a sign. And – what about the hamburger in the mouth? Is the chewing resistance, the aroma and the taste signs or is it maybe a real hamburger? Well, compare this to a blindfolded panel testing wine! With the wine in the mouth, a member of that panel might consider the perceived qualities of the wine to be signs of wine district, grapes and storage. In one case taste is not a sign, in another it is.

4. Is prosody a sign of emotion, or is it emotion? Is body language a sign of emotion, or is it the "real thing"? Are blushed cheeks signs of emotion or the real thing? Let us go deeper! Is high blood pleasure a sign or an emotion? Is the chemistry in our bodies, when we are aroused, signs or emotions? Well, to the doctor testing a depressed patient, these variables certainly are signs. Is then the subjec-

tive experience of these bodily states a sign of emotion, or emotion?

As we can see from these four examples, the existence of the sign depends on whom we ask and the situation. This decides the perspective. We cannot know if we are dealing with signs before we know the perspective.

It is easy to make the mistake to assume that Peirce's 1897 definition implies communication since it states that somebody is addressed. If there is an addressee we have to assume a sender and that this sender is intentional. But at a closer look at the definition we find that the addressee is addressed not by a sender but by the sign. The definition does not imply intention, not even a sender and thus not communication. In some cases of semiosis there is no sender. This means that the sign does not have to be encoded.

Let me exemplify this with the feather of the falcon. This indexical sign is not intended and not encoded. Can something not encoded be decoded? Somebody picks up this feather, examines it and draws the conclusion that this feather comes from a falcon. Is she decoding a sign? Yes, she is using semantic knowledge to decode the feather as a sign of a falcon. She conceptualizes the sign. A sign can thus be decoded even if it is not encoded.

Anything can thus be a sign, if it is culturally decoded. Even a walk in the forest is semiotic, if we choose that perspective. Trees and stones may remind us of people or goblins. There are signs of the age of trees, signs of season, signs of ecological mechanisms, signs of geology, signs of animals, signs of human activity etc. The nature as thus is a sign of God's existence, to some of us. But the forest is also a reality of its own, and is perceived and lived so by its inhabitants and occasional visitors. We interact with it. The paths shape our movements, the waters urge us to drink, the prey makes us hunt etc.

Sometimes we do reflect on what we perceive and sometimes we do not. Sometimes we act without hesitation and sometimes we think about the consequences. Sometimes the forest is just a background to a conversation and sometimes, as now, it is the focus of the conversation. Sometimes we use the interpreter and sometimes we do not.

Music can, just as the forest, be thought of and almost any association can be triggered by listening. It is important to stress that these associations may be emotionally charged and thus that they may affect us. Any element in music can be considered a sign. But music is not just semiotic; it is also lived by composers, performers, dancers and listeners in direct interaction.

The sign thus is perspective dependent. So are the concepts "code" and "representation". We say that music is encoded as a neural representation. This understanding is cultural. It is dependent on our knowledge in neurology. As perceiving subjects we are not aware of the neural code unless we take a cultural perspective.

Is music a language?

It is often said that music is a language. What exactly do we mean by that? Is music comparable to a national language? This question may sound naïve. To most readers it is clear that music is not symbolic the way national languages are. If it is not, what kind of language is it? I will make this comparison in a literary way. The purpose is to show that the eventuality of symbolisation in the Peircian sense can be ruled out.

What are the similarities between sounding music and spoken language? Language has one physical and one semiotic side. Louis Hjelmslev referred to this as two forms: the form of expression and the form of content (Hjelmslev 1953). The form of expression is also called emotional prosody. This is the melody (pitch change), the timbre (overtone constellation), changes in volume, and intonation aspects of language. Let me exemplify!

Let us say that the Prime Minister of Sweden is interviewed. Emotional prosody then is the perceptual difference between a reading a script of the speech and listening to an impersonator imitating the speech. A skilled impersonator will not only reveal information that you cannot get from the script (e.g. fear, bragging, uncertainty, personality) but ad his personal critique of the Prime Minister.

The form of expression and the form of content are linked to two different kinds of knowledge. The form of content is declarative and the form of expression is procedural. Prosody thus is learnt and communicated through imitation. The form of expression comes very close to music, and one of the ways to investigate music and emotion is to make this comparison (Juslin & Laukka 2003).

Is music semiotic? Yes, it is. As we have seen anything could be a sign. Music is no exception. But, in language the signs are Symbols – conventional arbitrary signs. Are there Symbols in music? No! Here are the reasons:

1. If music signs were Symbols, listeners would agree on the meaning of an instrumental piece of music. This could easily be tested. I have found no such investigation. The reason is of course that nobody truly believes that instrumental music can produce such shared meaning. It is from the beginning a hopeless enterprise. Music does not signify the same thing for you and me. A melody means one thing to you and another to others. The associations differ depending on who we are and the situation (Gabrielsson & Lindström Wik (2003).

2. In language it is not a fact that a change in temporality changes the information. You can read this sentence at any speed and read some words quick and some slow. The meaning would be the same. In music, changes in the note values would change the identity of the piece. The reason why language information is not affected by temporal changes is that it is symbolic. The symbol is totally abstract and does not contain a time dimension.

3. Contrary to language, it is hard to delimit and define a Symbol in music. Where does it start? Where does it end?

4. The smallest sign of meaning in language is the so called morpheme. This can, according to Hjelmslev, be defined through the commutation principle (Hjelmslev 1953): a morpheme is the smallest part that would change the meaning. To some extent this can be performed in music. If we for example insert a new base tone in a chord, the chord would change and so would the function of that chord. Another example is the change of a major third to a minor third in a

chord. But in language we can see that the prefix "un" changes real to "unreal". Negations are not possible in music. Musical signs do not follow logical rules. Symbols do, according to Peirce (1980–2000: 173)[348].

5. Since Symbols are arbitrary it is possible to translate one language to another. If music was a symbolic representation, it should be possible to translate music sign by sign to a national language. If so a verbal description of a piece of music would give the same meaning as listening to the music. This is obviously not the case.

6. Music semiotic theory implies that symbolical meaning has been established through the span of music history. But knowledge of music history is not needed to enjoy music.

I conclude that music is not a language in the same sense as national languages are. Language has one aspect in common with music – the physical aspect – the form of expression. But the most important thing, which makes a language a language – the culturally agreed code, it does not have. This means that there is no one to one relation between sound and content and thus music is just an "abracadabra". Play "abracadabra" it says in the score. It is all about the sound – the form of expression. Meaning and emotion is immanent in this sounding form.

From this follows that musical structures cannot be combined to narratives. Peter Kivy (2002: 182–201) argues that narratives in music is nonsense. There are simpler ways to tell a story than to compose a symphony, he states laconically.

Music can be interpreted as a sign or a set of signs, but everything can. Thus we can say that music "speaks" to us – metaphorically. But this is personal information. It depends on who we are and the situation.

Musicologists often present analysis under the pretext that the music signifies this or that, and it may, but since the signs are not symbolic, they are not culturally coded. What the musicologist is launch-

---

[348] Peirce, C.S. (1980–2000) Harvard lectures on the logic of science. *The Writings of Charles Sanders Peirce*. Vol. 1. (Ed. M. Fisch)

ing is therefore her personal associations packaged in semiotic terminology, which makes it sound more scientific and general than it is. We do not hear these signs until the musicologist has presented them and explained them to us. This may lead to a new way of listening. For other musicologists.

Can semiotic theory explain music emotion?

The position that music can be a sign of emotion is taken by Philip Tagg (Tagg 1979). This sign – *the museme* – is treated as a symbolical sign. It is argued that the emotional meaning of the museme has been established in music history and in music culture. Let me discuss this with a well-known example.

Let us recall the alarming music in the murder scene in Hitchcock's Psycho. Did the composer Bernard Herrmann use a sign of horror? Or is it horror lived by the audience – horror perceived? How did Herrmann know how horror sounds? Did he find this in his own body? Are we associating it with warning signals? Why then do warning signals sound like that? No one can misinterpret it. Or should we say misperceive? This sound seems to correspond to the bodily state of horror. It puts the body in an alarmed state. If this were a symbolic sign, any sound would do, since symbols (as words) are said to be arbitrary signs. This is hardly the case.

It is exactly this connection between music and the body that rules out that music could be arbitrary signs. Music is, as stated before, procedural knowledge once it is learned. Let me provide an example of a clash between semantic and procedural understanding of music.

Ian Cross has described his difficulties to learn to clap the tactus of a Bolivian rhythm (Cross 2003)[349]. The perception of tactus in Bolivia differs from most other Western music. Cross was able to learn how to clap this rhythm, but it felt theoretical; it did not appeal to feeling. This testimony indicates that he was trying to learn the rhythm intel-

[349] Cross, I. (2003) Music, cognition, culture and evolution. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre) Oxford University Press, New York. 52–56

lectually and succeeded, but this new knowledge clashed with his body knowledge.

The so-called *modular theory* suggests that special-purpose brain modules automatically and subcortically process acoustic stimuli, resulting in perceptions of emotion (Juslin & Laukka 2003). I would say that this is more than a theory. I have already presented many other references indicating that music engages the limbic/paralimbic system prior to neocortical activity.

Modules are domain specific (Peretz *et al.* 1998). The left temporal lobe is involved in decoding and interpretation, while the right takes care of melodic and prosodic features and their emotional aspects. Right-sided temporal brain damage thus is one of many neurological lesions affecting the production of emotional prosody. This can be noted in speech as a flattened monotone melody of speech with a slowed down variation of pitch. The prosody is dissociated from inner feelings (the patient recognizes this and is disturbed by it). When it comes to perception, it has been noted that the sensitiveness to emotional prosody, facial expressions and gestures are affected by right-sided brain damage.

Interestingly the right hemisphere is engaged in music listening amateurs, while both hemispheres are engaged in trained musicians. The explanation is that musically schooled listeners have names for the events in music (Fagius 2001: 71)[350].

Brain dissociation is a strong argument against any theory claiming that semiosis causes musical emotions. Since music and language are processed by different brain modules, it seems far fetched that music perception should involve decoding of Symbols. This is supported by findings reported by Isabel Peretz. Composers who have lost speech as a consequence of brain damage are still able to compose; and reversely, brain damage can cause loss of all musical ability but not speech (Peretz 2003).

In concordance with these findings it has been shown that ERP curves are different for unexpected words (N 400) than for unex-

---

[350] Fagius, J. (2001) *Hemisfärernas Musik*. Bo Ejeby förlag, Gothenburg

pected tones (P 600) (Besson & Schön 2003)[351]. According to the authors this difference points to different brain processing of language and melodies.

The neurological picture is complex. Since both music and language are auditory, sequenced and melodic; since both contain forms of expression, several brain areas are used both in music and language processing. But the results show interesting differences as well, indicating that music does not symbolize.

A convincing argument against the proposal that music signifies emotion are the many brain scanning studies previously referred to indicating separate pathways for the processing of music structure and the processing of musical emotions.

It may be interesting to note that Peirce never claimed that music could signify emotions. In the only statement he made about music he rather indicated the opposite: Musical emotions can signify. In a letter to Lady Welby 1904 Peirce exemplified the icon as:

> [a] vision, – or the sentiment excited by a piece of music considered as representing what the composer intended. (Peirce 1931–1958: vol. 8, § 335).

Let us say that a composer has the intention to illustrate the movements of the sea. Musical features similar to the movement and the sound of the sea would cause an emotional reaction in the listener. This emotional reaction would make her imagine the sea. The *feeling* (phenomenally perceived emotion) this piece induces in the listener then is an iconic sign of the sea provided that the feeling makes her think about the sea, according to Peirce.

This possibility is not at odds with the neurological findings. Rather there is anecdotal proof that musical emotions can signify. Recall the patient who suffered from amusia but was able to identify music titles from the emotion. It is however doubtful whether this emotion qualifies as a sign.

There is thus not much in favour of the idea that music could signify emotions. But, what about the Eva Cassidy example? I know

---

351 Besson, M. & Schön, D. (2003) Comparison between language and music. *The Biological Foundations of Music* (Eds. I. Peretz & R.J. Zatorre) Science and Medical Publications. Oxford University Press, Oxford. 267–293

that she died from cancer. When I first listened to her interpretation of 'Over the Rainbow', this came to mind. The song is an indexical sign of Eva Cassidy's fate. However an important aspect here is the fact that the words of the song are understood in the context of the singer's fate. The line "Somewhere over the rainbow, skies are free" means in this context that death is a liberator of the present pain. Not until this text dependant cultural connection is made, can the song signify this breathtaking longing for liberation. After just one listening occasion, the music becomes charged with this emotion and is thus perceived emotionally even if I do not listen to the lyrics or explicitly make the connection to the fate of Eva Cassidy. But this does not mean that the music signifies emotion.

Conclusions

1. The listener may be cognitively conscious about music, but she may also be just phenomenally conscious or just listen pre-attentively.

2. Any event in music may act as a sign.

3. There is no sign until a perception is culturally decoded. This has two consequences. A) The mere perception of a music feature does not make it a sign. B) The fact that somebody perceives a sign does not automatically make it a sign for somebody else. The sign is perspective dependent. The old watershed in music philosophy whether music is a representation or not must thus be answered politically: Yes and no. It depends.

4. Music is not a language, since music is generally not a system of Symbols.

5. Music can have emotional meaning to the listener without being a sign. I would call this view *enhanced formalism*. Music can be a form that does not re-*present*. The emotional meaning is *present* in the form.

If we should formulate these conclusions in one sentence, it would be: Signs exist only in a cultural (cognitively conscious) perspective, but since musical emotions are pre-reflective, semiotics is not the answer to the music emotion enigma.

7

Emotion

What is an emotion?

Although emotion research has a long history, scientists have not yet agreed on what they are investigating. This fact tells us something about how difficult this subject is. Some argue that they know exactly what an emotion is (Rolls 2005)[352] while others (Tomkins 1991)[353]; (LeDoux 1996)[354]; (Panksepp 1998) rather state that emotion is an umbrella concept for a variety of related phenomena. Furthermore some claim that they know where in the brain emotions occur while others insist that this is a perception action phenomenon engaging the whole body. Since the concept emotion is so evasive I have, up to now, focussed the discussion on perspective specific musical emotions.

A question that has become more and more acute along this investigation is: On what level of awareness do emotions emerge? As we have seen perceptions and emotions are to a certain extent processed along different pathways. How does emotion relate to perception? Does the perception cause the emotion or does the emotion cause the perception? Is emotion a perception?

It is time to devote a special chapter to the emotion. Let us start with a statement by Antonio Damasio:

---

[352] Rolls, E.T. (2005) *Emotion Explained*. Oxford University Press, Oxford

[353] Tomkins, S. (1991) *Affect, Imagery and Consciousness*. Springer Publishing Company, New York

[354] LeDoux, J. (1996) *The Emotional Brain* – the mysterious underpinnings of emotional life. Touchstone, New York

> All of the machinery of emotions is triggered and brought into action in an involuntary, non-conscious manner. So we really have no control over how the concert of an emotion is going to be played. (Damasio 2003)[355]

This involuntary and non-conscious nature of emotion may seem surprising since we can feel happy or sad and thus are seemingly aware of our emotions. Furthermore, culture demands that we exercise some control over our emotions.

What does Damasio mean? We use different parts of the brain to produce spontaneous smiles and fake smiles, he states. If you ask someone suffering from stroke to smile only half the face will smile. But if you tell a joke to the same person you will see a symmetrical smile. This observation could be connected with Merleau-Ponty's example of the man who could not put his hand on his head when asked to, and then suddenly smashed a fly that had landed there. These examples illustrate that voluntary actions following on cognitively conscious decisions differ neurologically from spontaneous procedural actions. As mentioned in Chapter 3, it has been found that facial mimicry reactions measured by EMG (electromyographic activity) are to fast to be conscious (Dimberg *et al.* 2000). Furthermore this first fast reaction can be replaced by a culturally more suitable expression. This has been called emotional regulation (Gross 2002)[356]. If, for example, the first reaction is aggressive, it is replaced milliseconds later by a smile. The subject is not aware of the first emotion, but of the second. Thus this first emotion fits Damasio's description. Some emotions belong to the involuntary domain. The biologist/neurologist understanding of emotion obviously differs from the everyday use of the word, which refers to the subjective valence of our present state – the feeling. The conscious observation of emotion is believed to be specific to humans and thus does not cover the biologist's use of the word. The every day understanding demands cultural development of the concept emotion. If there

---

355 Damasio A.R. (2003) Radio interview in Emotion and the Brain from *The Gray Matters* radio series produced for Public Radio International, in association with the Dana Alliance for Brain Initiatives.

356 Gross, J.J. (2002) Emotional regulation – affective, cognitive and social consequences. *Physiopsychology,* 39. 281–291

were no such developed concept, we would not be able to separate the emotion from the world perceived. The biologist understanding does not presume such a concept.

In the eyes of the biologist emotions are important steering mechanisms for animals, including human beings. Even in bees, we can see substances reminding of, and with the same function as, dopamine (Quartz & Sejnowski 2002). Do bees have emotions? Yes, in the biologist sense that bees have a system designed to steer behaviour, reminding of our emotional system.

Why do we have to care about this bee level understanding, when it is the human experience of emotion – the feeling – that really interests us? The answer is that in order to understand the feeling, we need to understand biological emotion. The feeling is after all a symptom of the emotion.

Other symptoms of emotion are behaviour, high blood pressure, blushing, substances in the blood, facial reactions etc. but the emotion itself eludes us. What then is a symptom of what? To find the emotion is like peeling an onion. There is no core. This is in agreement with Damasio's statement. We are not aware of the emotion. With a biological understanding of emotion, the symptoms are the only thing we can become aware of.

The lack of awareness and control make emotions difficult to research. It is not like the research of physics or even the research of cognition. Emotion has for periods been considered not researchable (LeDoux 1996).

The difficulty of pinpointing what an emotion is has caused debate among neurologists since the anatomy of the emotion circuits depends on how the concept emotion is understood. Most researchers consider the limbic system as the emotion circuit whereas LeDoux (1996) even denies that there is such a thing as a limbic system designed to process emotions. He argues that every emotional reaction has a function in a perception-action system, which involves the whole organism. It is dubious, LeDoux reasons, to appoint a specific brain area as the emotion circuit.

Although limbic activity coincides with emotional symptoms, the whole process of perception-action is emotional. Emotion directs perceptive attention (Vuilleumier 2005)[357], amplify cortical response to stimuli (Schupp *et al.* 2003)[358], motivate action, guide action and feeds back action. Every aspect of our lives is emotional and for this reason it is hard to deny that every perception-action process is emotional some way or other. I would go as far as to suggest that James Gibson's affordance (see Chapter 2) is the emotional reaction to the object perceived.

There is thus a vast diversity of phenomena and functions that are referred to as emotional. Just consider the different modalities! Are emotions elicited by sight, hearing, touch, taste, smell and bodily reactions equally? Does the feel of a soft surface have something in common with the pleasant aroma of a rose? Are the emotions of these experiences the same? Perhaps emotion is a generic term for a variety of phenomena. Probably it is a simplification to talk about one emotion circuit. Jaak Panksepp identifies seven circuits with different functions and connections: seeking, rage, fear, panic, lust, care and play (Panksepp 1998).

Of course these pre-attentive, basic and biological reactions can be connected with cognitive ideas and be attended to. But since music seems to affect us on a direct level, the understanding of the biologically unattended reaction is needed.

The neurophysiology of emotion

There are no unambiguous interpretations of brain scanning pictures of emotion processes. This is due both to the problems stated above (what are we looking for and is there a specific area for this) and the complexity of the anatomy of the emotional system, if we at all can speak about an "anatomy". The fact that a brain area is proved to be involved in an emotional reaction is just one piece of evidence. It takes further research to pinpoint the function of that area. It could

---

357 Vuilleumier, P. (2005) How brains beware: the neural mechanism of emotional attention. *Trends in Cognitive Sciences*. Vol. 9. No 12. 586–595

358 Schupp, H.T., Junghofer, M., Weike, A.I. & Hamm, A.O. (2003) Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *Neuroreport*. Vol. 14. No 8. 1107–1110

do the job, be inhibitory or just be a relay station. Areas reported to be involved in fear, such as the amygdala, are sometimes reported active in positive emotions as well. Different parts of the same area may have specific functions.

The auditory processing of cues for identification of source and emotional expression is important to our understanding of social situations. This requires sophisticated representation. A multitude of neurons is devoted to these tasks. Sensory information is projected into the hippocampus, the orbitofrontal cortex, and amygdala for emotional valence (Rolls 2005). Many other areas in the limbic system and the paralimbic system have been found active (for a review of the anatomical specification of limbic and paralimbic structures see Mega *et al.* 1997)[359].

The hippocampus generally maps for place, faces and episodic memories. It is involved in emotional valence. The orbitofrontal cortex monitors rules of the type – if this cue is present, then this reward or punishment will follow. It is the basis for pre-attentive decisions. It has been shown that orbitofrontal cortex is involved in rapid one-trial reversal of emotional behaviour. This means that it is engaged in switching the rule, or, in other words – in action/outcome learning (Rolls 2005). Damage to the orbitofrontal cortex leads to lack of emotional adjustment (uncooperativeness) and uncontrolled impulsiveness, but also impairment in the ability to process voice expression. Individuals with frontal lesions have shown unrestrained sensitivity to music. This indicates inhibitory function.

To say that the hippocampus recognizes the cue, prefrontal cortex applies the rule and amygdala initiates the emotional signals is a simplification. There is also feedback from the body informing the brain of the current state and the current needs. This information steers the perception.

---

[359] Mega, M., Cummings, J. Salloway, S. & Malloy, P. (1997) The limbic system: an anatomic, phylogenetic, and clinical perspective. *The Neuropsychiatry of Limbic and Subcortical Disorders* (Eds. S. Salloway, P. Malloy & J. Cummings). American Psychiatric Press, Washington DC. 3–18

The amygdala and the orbitofrontal cortex feed into the hypothalamus for autonomic/endocrine responses that influence the organs via the brainstem in order to prepare the body for action.

The amygdala and the orbitofrontal cortex also project into the striatum in the basal ganglia for implicit motor responses. The striatum is rich in transmitter substances, particularly dopamine, which activates movements.

Striatum feeds into the thalamus, which in turn projects into premotor supplementary motor and motor areas of the cortical brain.

While these commands are given, the neocortex may still be working on the conscious identification of the sound. When we become aware of the object causing the reaction, the reaction has already occurred.

Musical emotions

There is an abyss between neurophysiological research on musical emotions and the emotional experience of listening to Brahms. This abyss can never be bridged. These are two different perspectives.

Although the progress in neurophysiology is fast due to new techniques, the experiments are designed to answer the most basic questions. They are not on the Brahms level but rather on the major/minor level. And even this is not solved other than in the sense that this or that brain area is involved. This research can however help us see what is possible and what could not be. In the following I will present new reports as well as repeat some of the information given in earlier chapters in order to give a comprehensive picture of the neurophysiology of musical emotions.

As we have seen the hippocampus, the amygdala and the orbitofrontal cortex are generally activated in emotion processes. It has been found that the hippocampus and the amygdala are activated by music (Brown *et al.* 2004a; Levitin 2006). So is the orbitofrontal cortex (Blood *et al.* 1999; Koelsch *et al.* 2005; Menon & Levitin 2005).

Singing, listening, and imagining music generally activates auditory cortices, prefrontal areas (BA47) (Janata *et al.* 2003), primary motor cortices, the supplementary motor area (SMA), the insula, the cingulate sulcus, the thalamus, the cerebellum (Perry *et al.* 1999; Brown *et al.* 2004; Zarate & Zatorre 2005; Koelsch *et al.* 2005) and the mesolimbic system (the midbrain connection between the ventral tegmental area and nucleus accumbens) (Menon & Levitin 2005). The cerebellum is not, as was previously assumed, strictly motoric. There is increasing recognition that the cerebellum contributes to cognitive processing and emotional control (Schmahmann & Caplan 2006). Cerebellar lesions result in emotional disorders and impairments in just about any cognitive function. Cerebellum is linked to the limbic system and is implicated in imagery, and consequently in expectation.

Let us try to break down the sometimes contradictory and rapidly changing information to make an informed guess of the design of the music-perception-emotion circuit:

From its way from the ear to auditory cortex auditory information passes the colliculus at the brainstem. This area controls avoidance and orientation reactions. This is interesting because it is involved in anticipation and motor prediction. It contains maps of visual and auditory representation as well as representations of the body. These maps are "drawn" by the same neurons. In other words information from the senses unites here to construct a coherent representation of space (Berthoz 2000), which is related to the body schema. There is thus already at this stage a certain multimodal connection uniting heard information with space, and body movement.

In auditory cortex there are indications of two pathways: The ventral pathway for identification of source and one for movement detection.

The ventral pathway processes single sounds and single chords, as can be recalled from Chapter 2. Major, minor and dissonant chords cause reactions in the amygdala, the retrosplenial cortex (BA26), the cerebellum and the brainstem. Recent studies additionally show that the retrosplenial cortex is implicated in the perception of pleasant music (Brown 2004; Masaki *et al.* 2006[360]).

[360] Masaki, Y., Nishikawa, T., Ikejiri, Y., Tokunaga, H., Hatta, N., Uema, T., Kazui, H., Doronbekov, T.K., Ogino, A. Miyoshi, N., Tanii, H., Tanaka, T., Oku, N. & Takeda, M. (2006) Neural substrates of emotionally valenced episodic memory: A PET study using film stimuli. *Psychiatry and Clinical Neurosciences.* Vol. 60. No 1. 46–51

The dorsal pathway connects, as we have seen in Chapter 4, planum temporale with frontal cortices processing musical syntactic rules (frontal operculum) and rules of tonality (BA47). The planum temporale feeds into the superior insula, which activates pre-motor areas (Rolandic operculum) (Koelsch *et al.* 2006). We also know Chapter 3) that mirror neurons in pre-motor areas are activated by music. Koelsch's study indicates mirror neuron activity in the Rolandic operculum. Mirror neurons thus imitate music movement. Mirror neurons are connected with bodily representations in the brain (the body schema). The information is then projected back into the insula which is an area associated with the processing of feedback of bodily states. The insula feeds into the limbic system. Generally it has been implied that the limbic system feeds into the motor centres; with music we may have an example of the opposite. This supports the idea that emotions may be induced by imitated movement patterns. It also makes it possible to react emotionally to deviations from expected outcomes (music rules) since the rule can be compared with the imitated movement before the information reaches the limbic system. Probably this is a reciprocal relation. For instance, if we feel calm our breathing becomes slow, but we can also achieve calmness by slow breathing.

The pharmacology of music

There are about 150 transmitter substances acting on the synapses between nerves to affect neural activity. The very complexity of this system calls for careful interpretations. I just intend to sketch some possibilities relevant to music.

Several studies mention the striatum in the basal ganglia as an area activated by pleasurable music (Blood & Zatorre 2001; Koelsch *et al.* 2006). Striatum receives (along with the limbic system, and prefrontal cortex) noradrenalin (for drive and motivation), dopamine (motivation, movement) and serotonin (soothing).

Striatum is, as may be recalled from Chapter 4, involved in the sequence circuit. It organizes sequential movements. Dopamine is activating. It stimulates pre-motor cortex. If we cannot produce dopamine we cannot move (Parkinson's disease). As pointed out in Chapter 4, the internal clock used to time durations in the seconds-to-minutes range is linked to dopamine function in the basal ganglia (Meck 1996).

Dopamine is thought of as a motivation substance, since it is released when rewards are available (Schultz *et al.* 1997). Dopamine release also indicates uncertainty about the outcome – curiosity. As mentioned in Chapter 4 Jaak Panksepp connects dopamine with "the seeking system". Dopamine could thus be implicated in musically triggered movements and the stimulation of variation in music.

Dopamine is an active ingredient in the mesolimbic pathway (the connection between the nucleus accumbens and the ventral tegmental area). Opioid transmission in the nucleus accumbens is thought to be associated with dopamine release in the ventral tegmental area. As we have seen the mesolimbic pathway is activated by music listening (Menon & Levitin 2005).

A substance implicated in sequential movements triggered by the striatum is serotonin (Jacobs & Fornal 1997). This might explain why we love rhythmic music, swinging and dancing. It is possible that emotion in music, at least when it comes to the general pleasurable feeling can partially be explained by this role of the Striatum. The rhythmic quality of music thus seems to be one important factor to explain why we enjoy music.

The so-called chills (a sensation in the spinal chord associated with piloerection) in music engage the striatum, cingulate cortex and insula (Blood & Zatorre 2001). Chills are thought to be produced by the same reward/motivation system as in social bonding, parental love, loneliness and longing – and narcotic drugs like heroin and morphine (Panksepp 1995).

There are thus various examples of release of transmitter substances in striatum. This is interesting because of the drug culture connected to music. Known drugs are similar to, precursors to, or affect the level of, transmitter substances. Thus amphetamine, cocaine and nicotine increase the level of noradrenalin and dopamine. Serotonin is released by ecstasy, anti-depressive drugs such as Prozac, and to some extent by LSD and mescaline.

Music opens the taps of drug related substances in the brain. It is likely that people involved with music are seeking the musical creative stage by means of drugs. The natural high caused by the release of modest amounts of transmitter substances can be "helped" by artificial means, which intensifies the music experience. It may change the joy of dancing to rhythmic craze; chills to ecstatic love.

Causality between perspective and emotion

Does the perspective steer the emotion of the listener? To answer this question we need to be aware of the levels of awareness (cf. Chapter 2):

1. Pre-attentive levels: We have already shown that we act and react emotionally on this level. The perceived perspective thus does not affect such emotions. But, the choice of perspective is taken on a pre-attentive level. This implicit choice triggers a perception action process including the selection of relevant information. As we have seen each perspective actualizes specific information, which is processed by specific brain circuits connected to emotional areas of the brain. It is thus arguable that the pre-attentive choice of perspective precedes and causes emotional reactions.

2. Phenomenal conscious level: To deny that the perception on this level does not affect us is to say that we do not react emotionally to our conscious perceptual experiences. If so, I would not react emotionally when I become aware of the fact that some stranger at my door is armed.

3. Cognitively conscious level: To deny that this can affect us is to state that we do not react emotionally to what people tell us.

From this we can draw the conclusion that the evidence that emotions can be evoked on a pre-attentive level does not contradict that they can also be evoked on attentive levels. On level two and three the perspective clearly causes emotions (since it is implied in the perception). This said I want to direct the reader's attention to the fact that the enigma of music emotion is not to be solved on these two conscious levels. The problem is, as stated in the introduction that music often moves us although we are not aware of the reason. The emotion of interest is the emotion caused by pre-attentive perception processes. What does this imply for the perspective? The perspective is clearly a phenomenal experience. I suggest that perspectives can be pre-attentive and be processed in parallel. Otherwise we would not be able to switch between perspectives. On this level they may correspond to source domains. The most intense process is the winning hypothesis – the phenomenal perception.

Now let us reverse the question: Could the musically induced emotion change the phenomenal perspective? I suggest that this is the case. Although we can change perspective voluntarily, we normally do this automatically. Something steers attention and thus the phenomenal perspective. In the automatic choice this must be something pre-attentive. Musically evoked emotions could therefore be such a steering mechanism. Let me illustrate this with an example:

In a course of music in film I replaced the piano music added to a silent movie (Buster Keaton) with *Elegy to the Victims of Hiroshima* by Krzysztof Penderecki for pedagogic purposes. The result was amazing. The funny characters became bizarre and the comedy turned into an allegory about the absolute pointlessness of human life. This shows that the emotion induced by the music can change situation understanding. Even if I had an idea of what could be expected, I must say that the result astonished me.

Why did Penderecki's music have this effect? Keaton has the features of a tragicomic clown. The tragedy is connected to his appearance. It is in his face and in his movements. His movements are wrong somehow. They do not connect properly to the world. Humorous piano music does not help to identify with the clown, rather we identify with a laughing mob. The music sets the scene. It places us in a circus crowd. Every accident is accentuated and commented by the piano music. It is like the piano is saying: "Woops! Here we go again." We take a tribal perspective. Tragic music on the other hand, makes us take the perspective of the little man. We feel his helplessness. We take his perspective of the world as a hostile place. This is a dyadic perspective.

Musically evoked emotion may thus trigger the choice of phenomenal perspective.

To sum up, the pre-attentive choice of perspective causes emotion and emotion selects (among other agents) the phenomenal perspective. This is based on reasoning. It calls for further investigation.

Emotions and music – a hypothesis

The neurological evidence implies that music emotions exemplify the general involuntary and non-conscious character of emotions emphasized by Damasio. Music has been described as a *supraliminal affective stimulus* (Panksepp & Bernatzky 2002). This is evident in film music. Musically evoked emotions colour the characters. We do not have to be aware of the music for this to happen. The same effect can be seen when music is used as background in every day situations. We can create a setting for a dinner with music. Romantic music makes the other lovable. It makes yourself lovable, both in the eyes of the other and in your own eyes. If you played hardcore rock your partner would give another impression. And you would perceive yourself differently. The perceived other and the perceived self are emotionally coloured by the music. The emotion elicited by music can seemingly attach to any object.

If we are to take Damasio literally the emotion proper cannot be attended. But since music elicited emotions attach to objects attended, the emotion is experienced as a quality of that object.

Thus we can hypothesize:

1. Music evokes unattended emotion in the listener.
2. The object attended is coloured by that emotion.
3. The object attended and thus coloured by emotion can be the own body (feelings).
4. Or the music.

This hypothesis indirectly suggests a function of emotions. Objects are perceived as dangerous or attractive. There is an invitation/warning in the emotion. Emotions are not there to enrich our lives, although this is a nice side effect, they are there to help us survive. We do not think about our feelings when we confront something dangerous. Rather we experience the threat in the other and take action. We allocate our own emotion to the other where it becomes a quality – a feature, a characteristic. Sometimes this quality is symmetrical as when own love makes the other lovable and sometimes it is complementary as when own fear makes the other aggressive or when someone is crying and we want to comfort.

The prerequisite for this hypothesis is that we can focus any object as we listen to music and still be affected by the music. This demands that the emotion elicited by music and the perception of music are to some extent processed along separate pathways. The neurological evidence presented in the preceding chapters indicates that this is the case. As we have seen from the literature we are not aware of the emotion proper; and we do not have to be aware of the music to be affected emotionally. Even patients with amusia have emotional experiences of music.

The emotion produced by the music could be associated with any attended object. Not until we attend to that object do we become phenomenally aware of the emotion, but only as a quality of that object. The objects attended to could be objects in the environment, in fiction, situations, concepts, thoughts, memories or fantasies – anything we can focus. Interesting variants of such attended objects are the own body and – the music that elicited the emotion.

If the object attended is the own body, the body becomes emotionally perceived. This is the feeling: I feel tough, laid back, excited, good, sad, young, alive etc. It is a perceived quality of the body.

Now, let us say that the object attended to is the music at hand. This music evokes a reaction of sadness, say. If the music is attended to, our sadness clings to the music – becomes a quality of the music. The music sounds sad.

This hypothesis separates musically induced emotions from the definition of emotion presented by Edmund Rolls in *Emotion Ex-*

plained. According to Rolls emotions can be defined as "states elicited by rewards and punishers, that is, by instrumental reinforcers" (Rolls 2005: 11).

Music emotions are similar to, but still different from, emotions induced by a primary reinforcer such as a snake. Snakes elicit fear. Snakes are punishing. For this reason we avoid them automatically. We react emotionally to a snake before we identify it. The reaction makes us direct our attention to the snake. Then we can make the connection: "I hate snakes". The emotion becomes connected with an attended object. This is the likeness.

The emotion elicited by a reinforcer typically directs the attention to the reinforcer. But musical emotions do not do that. This is the difference.

Maybe it could be objected that the snake case is an over-explicit example, but even if we consider threatening sounds such as growls, we would identify the source. With music this does not always happen. Although no human being sounds like a symphony orchestra, we easily confuse this sound with the emotions of a character in a movie.

When the music in a thriller is scary, it does not take your attention from the situation in the film and make us identify the source. Rather it is the other way around. We cling even harder to the action in the film. Even if the music is alarming as in Hitchcock's *Psycho*, we still focus on the action. The stimulus music does not direct our attention to the stimulus the same way the stimulus growl make us attend to the dog. We would never confuse the growl with a friend walking at our side. We would not find the friend frightening. But this happens with music. Scary music makes film characters scary. There is absolutely no logic in it. But it is a fact. Music colours the world with emotions.

The point is: A reinforcer must connect the emotion elicited with the stimulus that caused the emotion. We must unmistakeably direct our attention to the stimulus that caused the emotion. Otherwise it could not reinforce behaviour directed to the stimulus. This does not happen with music. It is a mere coincidence if we connect the emotion with the music.

Music thus cannot be a reinforcer. It would be a disaster for an animal to fail to attend to a reinforcer. But it is not a disaster if we mistake music elicited emotions to be a characteristic in an actor. Music does not punish or reward behaviour the way reinforcers do.

Why is this important? Because Rolls claims that his definition of emotion is universal. Any exception from this definition would overthrow the definition, he says.

Maybe emotion elicited by music is the exception. This issue is however complicated because as we have seen, single sounds are reinforcers and music is made up by single sounds. But music is more than its sounds. If we heard a single musical sound, we would no doubt direct our attention to that sound. A possible explanation would be that music is an emotion communicated rather than a rewarding/punishing object.

The hypothesis presented reverses traditional attempts to explain why music moves us. The emotion does not have to emanate from awareness, identification, recognition, judgement, coding of signs or reflections on the impression of the music. Rather music produces emotions directly – subliminally. These emotions may, but need not, attach to the music that caused the emotion.

This is said with the reservation that we are now talking about primary emotions. To discuss the full palette of emotions elicited by music, a lot more must be considered such as all kinds of listener's perspectives and cognitive associations and the emotions connected to these associations. Here it is likely that music can act as a secondary reinforcer when connected to episodic memories (of home, childhood, a loved one etc.).

This hypothesis does not explain why music moves us, but it sheds some light on how music elicited emotions are recognized and how music animates the world.

The fact that the emotion elicited by music can cling to any object has made music the perfect tool for creating attitudes. Music has for this reason a widespread function in advertising, politics, religion and entertainment. Music forces attitudes on consumers to any displayed object.

The study of musical emotions leads to another important objection to established theory on emotion. As mentioned in Chapter 3 Darwin distinguished six categories of emotions derived from facial expressions. It has been argued that these expressions are universal, inherited and thus basic (Ekman 1994)[361]. The idea that we have a limited number of basic emotions has become cemented in psychological literature. I question this idea. The fact that the concept “basic emotions” is derived from reports depicting facial expressions alerts the suspicion that the proposed number possibly reflects that we have a limited capacity to phenomenally distinguish and cognitively name facial expressions. Furthermore, since musical emotions to a large extent are pre-attentive, we cannot tell how rich they are. My own phenomenal experience of musical emotions tells me that the limited number is absurd. I do not jump between six emotions as I listen to music. Rather it is a fact that music, as reported by some subjects in Gabrielsson/Lidströms Strong experiences in music, evokes experiences richer than words can express (Gabrielsson & Lindström Wik 2003). If music, as has been argued in Chapter 3, activates the sensory motor schema, this would lead to a continuous multidimensional space of emotions. Rather than six.

Conclusion

I have suggested a hypothesis. The premises are that emotion in a biological sense cannot be attended or controlled and that music affects us subliminally. The hypothesis states that emotions elicited by music colours what we attend. This includes the own body (feelings) and as a special and intriguing case: the music heard. This explains why music can be perceived as emotional. This hypothesis is not at odds with neurological findings.

[361] Ekman, P. (1994) All emotions are basic. *The Nature of Emotions. Fundamental Questions* (Eds. P. Ekman & R.J. Davidson). New York University Press. New York. 15–19

8

Conclusions and Discussion

Conclusions

I have applied *perception-action theory* to music perception. This brought us to a *perspective theory for music perception and emotion*.

Human behaviour is traditionally explained genetically or culturally. Musicology traditionally emphasizes culture in the analysis of music. Music listening is seen as a "language" and the encounter of new styles is analyzed as "enculturation". The perspective theory shows a third possibility: procedural knowledge achieved by imitation and perceptual learning. This is neither innate, nor cultural. It is often pre-attentive and for this reason easy to neglect. In order to make this level visible I have given culture a delimited definition connected to Ronald Posner's concept mental culture.

The egocentric, the dyadic, the allocentric, the tribal and the cultural perspectives have been discussed. Every perspective entails emotional implications.

In this conclusive chapter I will suggest an overall understanding of the key question: Why does music move us emotionally?

What is the function of musical emotions?

If we do not settle with the mere proposition that music triggers this or that brain circuit to react and cause emotion, we need to connect the emotion with use. To explain use is to give a reason why. What, we may ask, is the affordance? However it is hard to find a general use aspect of music listening (other than pleasure which would be a circular argument: We love music because we love it).

As may be recalled from the introduction, abstract or complicated behaviour – *target domains* – can be derived from concrete basic behaviours – *source domains*. If we can find the source domains of music listening, we can theoretically derive the emotion as a function in basic perception-action processes.

Is then music listening source domain processing? Music listening engages like no other activity the whole brain. We have seen that the auditory cortex, the frontal lobe, the pre-motor areas, the cerebellum, the midbrain, the brainstem, the basal ganglia, the corpus callosum, the insula, the cingulate cortex, the limbic, and the paralimbic areas are activated. Circuits, areas and pathways for sequence and rule processing, for movement-emotion, for imagery and for singing have been proposed in brain scanning studies of music perception. These are circuits on different levels connected in the overall perception-action processing of music. The variety of circuits involved supports the idea that music listening is not a specific and inherited process but a complex of basic processes.

We may ask: Are there no music specific areas? What about the processing of tonality, which has been traced to a specified area in prefrontal cortex?

The answer is that tonality is not universal and it is a relatively new organizing principle in the history of music. These facts undermine the idea that the perception of tonality is inherited. Tonality must be learned. The area processing tonality must be a general rule area.

What about the perception of pitch? As we have seen the processing of pitch involves not just auditory areas but also rule processing in the prefrontal cortex and cerebellum. There is thus no specific area for pitch processing. Separate areas must interact to produce the perception of pitch.

Each of the suggested perspectives entails use aspects. In the following I will suggest four principal source domains and connect these to the perspectives. The cultural perspective does not qualify as a source domain, since it is complex and abstract. For these reasons the cultural perspective will not be taken into account here.

According to metaphor theory the source domain can be actualized by the target domain as a metaphor. The persona (the other) and the landscape are such metaphors for music.

On a pre-attentive level the target domain “music” may be processed by source domain circuits. It could thus be the case that we process music *as if* it was an expressive person, or *as if* it was a landscape; that we use circuits originally “designed” for the perception of the other and the landscape, and that these processes generate emotions. The “*as if*” aspect was treated in Chapter 1, where it was connected to imagery. If we process music *as if* it was something else, this process may activate memories associated with the source domain and thus the source domain could appear as imagery.

I suggest four source domains underpinning the target domain music: the perception of single sounds, the implicit understanding of the other, finding the way in the environment, and joint action. Each domain is a perception-action process with a concrete use value and with domain specific emotions. Since these processes entail a selection of information, I suggest that these *source domain processes are pre-attentive perspectives*.

The perspective theory of music and emotion

1. **The Egocentric Perspective**: The landmark is the own body (the body schema). The knowledge this perspective uses is the automated knowledge how to react to primary and secondary reinforcers.

   Source domain: The identification and localization of sound source leading to immediate reactions. We are especially sensitive to voices. Calm voices are even and sonorous; stressed voices higher in pitch, loud and yell; dominating persons are loud and submissive persons have weak voices etcetera. The voice quality reveals the identity of the speaker. Unexpected sounds are generally frightening as are unexpected sound qualities and loud sounds. The egocentric emotional response is a pre-attentive reaction to a reinforcer and the function is to steer the action of the listener vis-à-vis the sound source.

Relevance to target domain: Music consists of sounds: timbre of the instrument, timbre of instrumental constellations, volume, pitch, and chords (consonance, dissonance major/minor) are single sound parameters. We react emotionally to single musical sounds and most of all to voices or instruments that are close to voice in timbre. As argued (Chapter 2) minor chords are more dissonant than major chords – perhaps not so much because the minor third is somewhat more dissonant than the major third but because of the clash between the minor third and the major third overtone emanating from the key note. Sound is important to musicians and listeners but is all the same sadly overlooked in the literature of music and emotion.

2. **The Dyadic Perspective:** The landmark is the other. This perspective selects information from procedural knowledge.

   Source domain: Understanding the situation from the other's perspective and understanding how to contribute in interaction. This implies a sense of the intention of the other. Through implicit imitation the observer copies the movement of the other to her own body or adds complementary action. This source domain is important for implicit interaction, communication, learning, and bonding. This is a pre-attentive mirror-neuron process activating pre-motor areas. These areas, which are connected to the limbic system via the insula, set off the pharmacology of emotion. If the pre-motor pattern is imitated the emotion evoked in the beholder will be the same as in the other. If, on the other hand complementary action is triggered, this pre-motor pattern will cause complementary emotions.

   Relevance to target domain: Music is, among other things, the emotional expression of the composer/musician. This expression has a movement pattern and a sound quality. There is an increasing amount of indications that this is latently and overtly imitated by the listener. This entails that dyadic emotions are pre-attentively communicated. Complementary emotions have not been tested for music. The fact that the mirror neuron system is more active in complementary action than in pure imitation opens up the possibility that sad music could cause tenderness and love in the listener.

3. **The Allocentric Perspective:** The landmark is in the landscape/structure. This perspective combines place, rule and procedural knowledge.

   Source domain: Finding the way. This perception action system engages the frontal lobe (rules processing), hippocampal structures (mapping of place in a structure) and cerebellum and basal ganglia (sequence processing). Implicit rule and sequence processing create pre-attentive expectations. There are three implications: Anticipation of structure, the "seeking system", and technique. We anticipate the structure. Mismatches between perceived actual outcome and predicted outcomes, cause emotional reactions. This surprise may depend on a clash between anticipated structure steered movement and imitated movement. It alerts the animal. If there is a mismatch, something unpredicted is going on. This emotion steers the attention to whatever that may be. Jaak Panksepp's "seeking system" connects structure with body movement. The drive of finding (curiosity) is dopamine regulated and this substance activates seeking behaviour. This is triggered by the possibility of reward. It activates body movement.

   Another connection between structure and movement is the technique. Technique is about finding the movement adapted to the structure with a minimal cost of energy. Since economizing with energy is vital to all animals, we can suspect that there must be an emotional reward/punishment feedback system for technical perfection.

As we have seen accuracy in monitoring activates the reward system.

Relevance for target domain: Most music has allocentric landmarks. These landmarks are a prerequisite for the perception of music as a structure following musical rules. Places in this structure can be anticipated. This helps us entrain the flow of music and to feel the deviances from expected continuances. Such planned deviances are powerful tools for composers.

When it comes to "the seeking system" the connection to music is perhaps less obvious, but we know that movement and dopamine production are connected. It is hard to believe that music makes us move because of curiosity. But music is engaging if it contains unexpected events. Surprise reactions could affect the production of dopamine and make us more attentive. Even if we move because of mirror neuron reactions, the dopamine production may accentuate those movements.

Technique is of major concern for musicians, as it is for listeners. Accurate technique is felt through effortless timing, pitch, sound and phrasing. Just consider the difference between a performance distinguished by effortless timing, pitch, and accuracy and the opposite: struggling to hit the right tone and follow the rhythm! We feel the effortless performance as talent – as grace even. Bad technique can turn beautiful music to horror. Technique is an overlooked source of pleasure in the literature of music and emotion.

4. **The Tribal Perspective:** The viewpoint is taken from the perspective of the interacting group. This perspective, which concerns our behaviour as group members, depends on procedural and episodic memory.

   Source domain: Joint action – collective interaction to achieve common goals. Joint action has been explained by mirror neuron theory. A shared perception of the situation is created. This process allows us to implicitly understand the goal of an interacting group and tells us how to contribute. If the joint action is a ritual, this will trigger episodic memories, from earlier rituals (since the memories are connected to the movement patterns). These memories may be emotional (as when the national anthem is played in sport ceremonies and the whole stadium is united in national pride). Tribal emotion unites the tribe in action towards a shared goal (e.g. going to war).

Relevance to target domain: The mere bodily synchronization with others in repetitive rhythmic movement is in itself pleasurable. Music is used in rituals and is in itself ritualistic in the sense that an established behaviour pattern is reactivated. Traditionally, music is an element in the ritual. The ritual creates a common understanding of the world – a mythology. It actualizes episodic memories from former rituals. The emotion is connected to this collective understanding (as when we sing Christmas carols). Playing music in a group is the ultimate example of joint action. The audience gets involved in this interaction. Even when we listen to music alone we interact with the performing group. We have imitation tendencies and we have tendencies to contribute – to fill in and complete.

Music is sometimes been described as abstract – as a language even. From this viewpoint use and emotion are puzzling. But if we consider music as a ritual, it becomes much less abstract. The affordance becomes visible: Come and join in! Dance with us! Sing with us! You may sit at a birthday party; you may eat the cake, but nothing really happens. Then a musician comes in. She sits down at the piano and finds the key songs that unite the party. We sing along and now it happens: This is us! This is the love we share! To sing and dance with other people is to be embraced by them.

Collective mourning at funerals is a universal phenomenon. Sadness is the emotion that is most easy to identify in prosody and music, maybe because we hear crying and sobbing. Crying is highly contagious. If I were to point out one powerful emotion in music it would be this: the emo-

> tion of longing. It is the experience of how music tears your heart out of your chest that really makes you wonder about the inexplicable power of music. The other side of this coin is the possibility of complementary emotion. If the music can express sadness, and melancholy, it could trigger the complementary emotion love and tenderness.

This theory suggests a relation between the perspective and the emotion. But what is the causal connection? Are all four source-domain processes simultaneously activated as we listen to music? I suggest they are, to some degree, and that the difference in intensity decides which perspective will present itself on the screen of consciousness. This means that we have many candidates for the emotional reaction to music. Can they all be responsible?

The case of Clive Wearing (Sacks 2007)[362] is interesting in this context. Mr. Wearing is an eminent English musician. In 1985 he was struck by *herpes simplex encephalitis* affecting the hippocampus. Wearing completely lost his ability to encode memories and thus developed total amnesia. Thus he has just moment-to-moment consciousness. But his musical skills are intact, with one important exception. If the music stops, he cannot proceed.

According to Sacks, the explanation is that Wearing has a procedural memory of music. Procedural knowledge is implied in the dyadic, the allocentric and the tribal perspectives. With the reservation that egocentric reactions are not learned by means of imitation, procedural knowledge could even be claimed to be involved in the egocentric perspective since the automated reaction of sounds is a bodily and implicit reaction. The involvement of procedural knowledge is the important difference between these four perspectives and the cultural perspective. It makes music embodied. It explains three typically musical phenomena.

1. The fact that the procedural memory is based in subcortical areas (cerebellum and basal ganglia) explains why this memory is robust. Music can be remembered by old people, when other memories are gone. It can be remem-

[362] Sacks, O. (2007) The Abyss – music and amnesia. *The New Yorker*. Sept. 24.

bered by patients suffering from stroke when speech is lost. We may forget declarative facts but we will not forget how to move.

2. It explains why music can only be remembered by piece by piece.
3. And, what is most important for the music emotion question, this musical information can be communicated by means of imitation.

The answer to the music emotion enigma is linked to the fact that music is procedural knowledge. The practice of music is procedural. Playing, listening and music imagery is procedural. Let me illustrate the importance of this by pointing out the difference in intensity between the emotion when your country is named, and the emotion when national anthem is played and sung. In the last case you are involved in a procedure, which engages the body. This movement pattern evokes episodic memories. Likewise, there is a difference in emotion when you read a poem you know by heart and new prose, even if the content is the same. Some people simply cannot read poems, because the voice will break. It is the developing schema (the temporal topography), which actualizes movement-emotion connections in the listener. The connection between movement and emotion is twofold: The movement is an emotional expression and/or it evokes episodic memories.

Let us re-synthesize the processes by connecting the source domain processes to the target domain music in order to get a comprehensive picture:

As we listen to music we process it as the emotional expression of the musician/composer. Since music is sequenced and rule based, we can anticipate the structure. These rules are to some extent style specific, which explains why unknown styles do not move us. By way of mirror neurons the physical expression of emotion is copied to the body of the listener. These expressions are the same as, or complementary to, the expressions of the musicians. In a concert the musicians and the audience entrain the same movement/structure pattern. Just as we follow performers in athletics, we follow the musicians as they struggle with the technique to hit the notes and to produce the sound. We take part as back seat drivers, and we feel the perfection in our own bodies. Everybody shares the emotional expression. A tribal mass effect of shared emotion and situation understanding is created, including emotional episodic memories connected to the movement pattern. Surprises in the music are mutually shared; emotions derived from sound shifts are mutually shared, the experience of technique is mutually shared. We all follow the same sensory-motor track on-line with the musicians. This sharing of the experience makes the emotion stronger. Even when we listen alone a joint action ritual takes place since we interact with the performing group.

Basically, music listening then is an emotional movement pattern communicated to the body of the listener. Listening to music you are connected bodily with others and share their emotional experiences. Even if there is a time gap between the production of a phonogram and the listening occasion, you interact with the composer and the musicians as if connected online. Thus, listening to music, you are never alone.

The fact that most musical emotions are pre-attentive paves the way for the hypothesis that these emotions can attach to any object that we happen to attend to. This makes music a powerful instrument to sell products, ideologies, and pop stars. Intriguing special cases are the own body (feelings) and music. This is thus the reason why we state that music can be felt as emotional.

This description is admittedly rough. It has to be because the research of mirror neurons and joint action is in a virginal state of development. Furthermore this is the first attempt to connect perspective with emotion in general and with music emotions specifically. There is thus no research on this at the present state and no hard evidence.

Discussion

The perspective theory has bearing on the discussion in music philosophy presented in the introduction. Since sounding music does not symbolize in the Peircian sense, it is not a form of content, not declarative, and thus not communicated on a cognitively conscious (cultural) level. Rather it is a form of expression. It is procedural and communicated on a phenomenal and/or pre-attentive level. This means that Hanslick was right when he stated that music principally is tonally moving forms. But he was wrong when he concluded that music cannot communicate emotions. As discussed in Chapter 2, form implies implicit meaning. This can be connected with Peter Kivy's concept *enhanced formalism*. The temporal topography of music has meaning. Contrary to Kivy, I claim that we are emotionally affected by this form. We do not have to be phenomenally aware of the form for this to happen. As with all procedural knowledge the form of expression is communicated by means of imitation. It has been shown that this imitation affects the listener emotionally.

Arousalists claim that music cannot have emotions since it is not a sentient being. For this reason they insist that music must be perceived as emotional because we project our own emotions into the music. This brings us to the objection against arousalism theory: "Arousal theory does not explain why different kinds of music elicit different emotions". This is true, but perspective theory does. Music causes pre-attentive emotions. This is the emotion we project into the music, where it becomes attended as a quality of the music. The contradiction between the proposition that we project our own emotions into the music and the proposition that these emotions are caused by the music is thus avoided. Both statements are true but from two different perspectives.

Another issue in music philosophy is: Why do we listen to sad music? I can see two alternative explanations:

1. Sadness is the easiest emotion to identify in prosody as well as in music. There is an endless line of music styles that express sadness due to loss: Emigrant songs, seamen's songs, love songs, country etcetera. Experience of loss is universal. It is closely connected to vocal expression and bodily expression. Thus it ought to be imitable. As we have seen imitation leads to positive emotions. This means that although we are imitating a sad expression and conjure this sadness, there is an embracing factor in the fact that we imitate. Empathy for the sad person is not an en-

tirely negative feeling. Empathy connects us. Empathy makes us warm. It is a prerequisite for love.

2. The finding that mirror neurons are activated for complementary action, paves the way for complementary emotions.

These two mechanisms probably interact. If someone dies, and your friend is devastated, her emotion from her loss is different from your empathic emotion. You feel what she feels but you also feel love. Musical emotions are of this second kind. They are thus not reactions to events but interpersonally communicated emotions. The implication is that musical emotions differ from other emotions because there is a positive factor in all musical emotions.

The perspective theory may also have some relevance for the developmental explanation of the music phenomenon.

There are two positions:

1. A position claiming that music is an inherited capacity. It has been proposed that music maintains large groups and that this indirectly would be beneficial for the spread of genes. It has also been suggested that music skill is an advantage in the struggle for mating partners.

2. The "source domain position" claiming that since there is no obvious use value in music, there cannot be any developmental advantage in it. Music is a composite of source domains, each with a developmental history of its own.

I have already shown (Chapter 5) that the "large group argument" could be derived from the tribal perspective; that it is the joint action aspect of music as a rite that creates emotional bonds and shared understanding. The large group advantage is thus not necessarily an argument for position number one.

When it comes to the "mating partner argument", I suggest two arguments following from the perspective theory:

1. On the dyadic level: An artist could implant implicit emotions in a listener that are perceived as attractive personal qualities of the artist. Powerful music makes the artist powerful, sexy music makes her sexy, sensual music makes her sensual and sad music gives the impression that she is longing for love and comfort.
2. On the tribal level: Consider the rites of fertility. They are universal; they are ancient and they can be seen on MTV. We can trace show-off behaviour in mating rites to the animal world. Anyone understands the invitation of moving hips in samba but cultural restraints prevent us from acting out.

This dissertation argues for a primacy of emotion. Perception is not only coloured by emotion but even steered by emotion. Thus we may ask: How has emotion affected the development of music? I propose that musicians must have felt: "This works and this works, but not this. If I do this, I communicate emotionally with the tribe." For this reason the development of music could be described as a long trial and error search for emotional response. Some usable sound structures – some *memes* (reproduced units of behaviour) (Dawkins)[363] – have crystallized. Other musicians have reused these memes in new compositions, varied them and developed them. As composers and musicians we see the effect in the audience, and above all, we feel the effect in ourselves. The emotion leads the composer and has thus been leading the development of music (together with other factors such as the development of instruments and cognitive knowledge). Due to geographical obstacles and cultural barriers these memes have remained local. The music we have today, the diversity of music all over the world, is to a large extent the result of this long random search.

---

[363] Dawkins, R. (2006) *The Selfish Gene*. Oxford University Press, Oxford. Chapter 11

Listening attitudes

The perspective theory has implications for strong emotions in music listening. The cultural perspective is in many ways an exception. Cognitive awareness takes the focus away from the phenomenal experience of music. This conclusion may seem amazing since it entails that the more we think about what we are hearing, the less we will experience on a bodily and emotional level.

So-called *deep listening* implies an altered state of awareness (Becker 2004). It is a state reminding of meditation. In meditation something concrete is focussed. This could be the own breathing, a body part, a drum rhythm or a mantra repeated over and over. Through this technique, cognitive awareness is cut off. A here and now state of phenomenal consciousness is reached. It is a dreamlike state reminding of hypnoses – a state of raised suggestibility. It is the state in which we are transformed into other beings in religious rites or guided by shamans on a journey. It is a state where the world is perceived holistically; a state where the self is dissolved in the world. It is a state of heightened sensibility, a state where we merge with the music, become one with the music. All these effects are evident in the investigation by Gabrielsson/Lindström on strong experiences of music. Any perspective could be taken; with the exception of the cultural perspective leading to cognitive awareness. The more we intellectualize the perception, the more we lose the experience.

I would also count the focussed kind of listening that emerges from perceptual surprises to the deep listening category. In Chapter 4 we discussed the seeking system. I exemplified with Pat Metheny's *The way up* how a composer can create perspective shifts that are surprising and felt in the body of the listener. The listener experiences flow – a totally consuming presence in the music.

The common factor in these two listening attitudes is the present moment perception, in the meditation case through a rejection of thoughts and in the flow case through the focus on present events in the music. This too is a consequence of the procedural character of music.

Future research

1. Philosophy: The perspective theory is not limited to the analysis of music. We need to identify the landmark and the frame of reference for a full understanding of any situation involving perception, interaction, communication, signification and emotion. If the suggested theory is accepted, a general philosophy of perspectives is waiting to be carried out.

2. General application: Since learning can be pre-attentive, we should ask how entertainment violence in sports, media and computer games affect us. And since this learning is established on an implicit level, how do we deal with violent tendencies?

3. Cross modal effects: I would like to see an investigation where the principles for musical sequences and rules were applied on moving abstract pictures to see if this would in anyway cause musical experiences.

4. Neuromusicology: I have listed five experiments that could be carried out to answer the questions we have today in the discussion of the dyadic perspective.

5. How does the perception of technique affect musical emotion.

Finally I want to expand on a music-psychological issue. We know that any style or genre could induce strong emotions. It is, according to Gabrielsson/Lindström rather the fit between the situation, the listener and the music that matters. The exploration of this fit is a challenge for musicology. We have seen indications that pre-motor imitation produces pleasurable experiences. This may guide us to an understanding of the fit between the listener and the music. If the music is imitable, it will evoke positive reactions. I suggest that this does not exclude sad music, since we obviously like to listen to it.

What then makes music imitable? As we have seen in Chapter 3 familiarity with the object is a general condition for imitation. I speculate that we must have a memory of the musical pattern – a sensory motor schema. Another and complementary possibility is a familiarity with the emotional expression. This does not have to be musically imprinted. The listener could have an emotional experience, which makes her sensitive to the musical expression of that emotion. She could have a personal disposition for this emotional expression. Such sensory-motor patterns may have been established in childhood, due to events of strong emotional valence. Let me provide an example.

In *Solei Noir* Julia Kristeva explains melancholia as a bodily and painful memory of a mother lost in infancy (Kristeva 1989)[364]. Since the loss is pre-linguistic there is no name or symbol to attach to the bodily felt suffering. The memory of the mother (the sun) is black. Generally *black suns* then would be embodied memories of pain from infancy not accessible to cognitive consciousness. This causes the melancholic to phrase and rephrase her pain in vain, never being able to accurately formulate her loss. Kristeva analyzes several pieces of art and poetry as consequences of this drive to formulate the loss. The artist falls in love with the beauty of the artistic expression, which finally may become a substitution for the lost mother. The loss can be detected, Kristeva claims, in *semanalysis* which is an analysis of prosodic traces from the pre-linguistic phase. The existence of a "black sun" thus causes a pain that cannot be communicated as a form of content, but it can be detected as a form of expression – as an oceanic longing. This is an example of a purely emotional memory – a memory that is not attached to an object. It is thus implicit. Such memories leave prosodic traces (Charles 2005)[365]. An explanation of this phenomenon is of course the connection between primary emotions and action. This memory is procedural since it is bodily and leaves traces of emotional expression.

If, then, the listener has the same emotional experience as the artist, she would have the same sensory-motor pattern as the artist – and the expression of the artist would be imitable. She would implicitly recognise the cry for the missing mother in her own voice.

---

364 Kristeva, J. (1989) *Black Sun*. Colombia University Press, New York

365 Charles, M. (2005) Patterns: Basic unites of emotional memory. *Psychoanalytic Inquiry,* 25. 484–505

Final words

It was my emotional reaction on the Brazilian singer Gonzaguinha that made me think about music and emotion. I could not understand how this singer from a totally different culture, singing in a language I did not understand, could move me so much. I do not claim to have the answer now, but I can see that there is a match between Gonzaguinha and myself. This might cause imitation. We were both songwriters and guitarists. He was playing Brazilian music and I had a few bossas on my repertoire. We were both dwelling in the zone between folk songs and jazz. We were the same age and to some extent expressing the same concern for the issues of the era. We were young in the political 60s. We were both university kids. This might sound cognitive but I believe that it is to a considerable extent embodied. To be a radical at the university was to some extent to belong to a tribe, sharing the attitudes, the type of clothing, the haircut, the music and even the movement pattern. This is something, but it is not enough.

The summer I heard Gonzaguinha was the summer after my father had died. My mother had passed away a few years earlier. It was thus the first summer without my parents. Everything, the path we went to fetch milk from the nearby farm, the path to the sea, the very stones on these paths called them back.

I was particularly taken by Gonzaguinha's song *Fotografia* (photo). There was this drone base producing dread and, at the same time, an oceanic horizon. Over this, the sensitive voice with outbursts of falling lamenting wordless singing.

The song inspired me to do my own version called *Sorrow*. My song was about the paths where I used to walk with my parents. I later learned that Gonzaguinha's song was commenting on photos in the newspaper of the killings of street children in Rio de Janeiro in the early 90s. Gonzaguinha identified with these kids. He called himself a *moleque* – a street kid. One photo, he states in the song, shows a kid looking into the gun of his killer. This is an image of the loss of innocence.

The common feature of the two songs, the deep theme, was thus the loss of childhood. And the experience that inspired these two songs was in both cases an event that actualized the loss of childhood. Why then is the loss of childhood traumatic to both of us?

As it turns out, we have a traumatic childhood experience in common. We both have "black suns". Gonzaguinha lost his mother at the age of four months. You can, when you know it, trace this in his lyrics. The cover of the album Luiz Gonzaga Jr. shows how he could perceive himself: divided, broken, and smashed into pieces. There is a black hole in his forehead (Gonzaguinha 1973)[366].

I was, due to illness, separated from my mother for a couple of months at the same age.

* * *

The emotion elicited by music has a quality of fundamental truth. The experience is undeniable. Something…happened here. The very fact that there is nothing except the music itself that caused this experience makes us think that music moves us in very profound ways; makes us feel that this truth cannot be formulated. This makes the claim of truth stronger, because this truth is general – it is about nothing and everything. It is about life. There are no words for it…

You claim that this must be spiritual some way. That it is greater than us, ungraspable, God given; and speaking directly to our souls.

I can't deny that. It is beyond the scope of science.

[366] Gonzaguinha (1973) CD: Luiz Gonzaga Jr. EMI Brasil.

Sources

Personal information

Hultqvist, A. (2007) 20/4

Koelsch, S. (2007) 7/7

Moraes, J. (1999) 21/5

Eriksson, G. (2008) 12/1

Radio/TV

Bosco, J. (1990) Interview in Bar Brasil, a Swedish Radio series on Brazilian music by Lennart Peck, Lennart Kjörling and Sven Kristersson, part one. July 2

Buck, L. (2004) Nobel Minds, TV-interview. SVT/BBC

Damasio, A.R. (2003) Radio interview in Emotion and the Brain from *The Gray Matters* radio series produced for Public Radio International, in association with the Dana Alliance for Brain Initiatives.

Gonzaguinha (1979) interview by Radio Globo, April 2

Olsson, L. (1977) *Samba*, TV documentary produced by Lennart Olsson for TV 1 Sweden. March 2

Tramo, M. & Coleman, S. (2003) interviewed in Emotion and the Brain in *The Gray Matters* radio series is produced for Public Radio International, in association with the Dana Alliance for Brain Initiatives. www.mega.nu/ampp/danamusic

Recordings

Cassidy, E. (1998) CD: Over the Rainbow. *Songbird.* Blixt Street

Gonzaguinha (1973) CD: *Luiz Gonzaga Jr.* EMI Brasil

Gonzaguinha (1980) CD: *De Volta au Comeco,* EMI Brasil

Gonzaguinha (1993) CD: *Cavaleiro Solitario*, Som Livre

Gonzaguinha (1994) CD: É. *Mestres da MPB*, Warner

Insectoid (1997) CD: *Sacred Sites.* Quote from the folder of the CD. Return to the Source

Mozart, W.A. CD: *Piano Variations.* Vol. 3. Naxos

Metheny, P. (2005) CD: *The way up,* Nonsuchrecords

Internet

Boeree, C.G. (2006) www.ship.edu/~cgboeree/genpsyperception.html

Daw, N.D. (2003) Reinforcement learning models of the dopamine system and their behavioral implications. www.gatsby.ucl.ac.uk/~daw/thesis.pdf

Davidson, R. (2005) www.eurekalert.org/pub_releases /2005–08/uow-sbs082605.php

Janata, P. (2001) Neurophysiological mechanisms underlying auditory image formation in music. http://atonal.ucdavis.edu/manuscripts/Janata_2001c.pdf

Klapuri, A. (2001) Top Down Processing. www.cs.tut.fi/~klap/iiro/topdown.html

Levitin, D.J. (2007) www.iht.com/articles/2007/01/03/features/ music.php?

Meidell, K., Gaab, N., Halpern, A. & Schlaug, G. Neural correlates of performance in a motor imagery task: an fMRI study with professional pianists. http://208.164.121.55/hbm2003/abstract/abstract1086.htm

Port, R., Tajima, K., & Cummins, F. (1996) Self-entrainment in animal behaviour and human speech. *Online Proceedings of the 1996 Midwest Artificial Intelligence and Cognitive Science Conference.* www.cs.indiana.edu/event/maics96/ Proceedings/Port/port/.html

Rizzolatti, G. interviewed by Maddalena Fabri Destro on the net. http://scholarpedia.org/article/talk:mirror_neurons

Pärson, A. (2005) Interview in *Göteborgs-Posten.* www.gp/jsp/crosslink

Winkler, I. & Näätänen, R. (2006) Theoretical framework of pre-attentive auditory processing. www.cbru.helsinki.fi/ research/theory/description.html

Bibliography

Allwood, J. (2000) Activity based pragmatics. *Abduction, Belief and Context in Dialogue: studies in computational pragmatics* (Eds. H. Bunt & B. Black) John Benjamins, Amsterdam. 47–80

Alperson, P. (2004) Music: formalism and beyond. *The Blackwell Guide to Aesthetics*. (Ed. P. Kivy) Blackwell Publ., Malden, MA. 254–275

Altenmüller, E. (2006) Emotions in music are mainly learned. Paper presented at the conference *Music in the Brain* at the University of Aarhus 21–22 April

Arnold, M. (1960) *Emotion and Personality*. Columbia University Press, New York

Bach, C.P.E. (1949) *Essay on the True Art of Playing Keyboard Instruments*. W. W. Norton & Co., Inc., New York

Bach-y-Rita, P., Collins, C.C., Saunders, F., White, R. & Scadden, L. (1969) Visual substitution by tactile image projection. *Nature*. Vol 221. No. 5184. 963–964

Bakhtin, M. (1984) *Rabelais and His World*. Indiana University Press, Bloomington

Baldissera, F., Cavallari, P., Craighero, L. & Fadiga, L. (2001). Modulation of spinal excitability during observation of hand actions in humans. *European Journal of Neuroscience*. Vol. 13. 190–194

Bartels, A. & Zeki, S. (2000) The neural basis of romantic love. *NeuroReport*. Vol. 11. No 17. 3829–3834

Bayliss, A.R. and Tipper, S.P. (2005) Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology*. Vol. 96. No 1. 95–114

Becker, J. (2001) Anthoprological perspectives on music and emotion. In *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press, Oxford. 135–161

Becker, J. (2004) *Deep Listeners – music emotion and trancing*. Indiana University Press, Indianapolis

Bergson, H. (1910) *Time and Free Will: An essay on the immediate data on consciousness*. George Allan and Unwin, London

Bermudez, P. & Zatorre R.J. (2005) Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences*. Vol.1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo). 395–399

Berthoz, A. (2000) *The Brain's Sense of Movement*. Harvard University Press, Cambridge, Massachusetts

Besson, M. & Schön, D. (2003) Comparison between language and music. *The Biological Foundations of Music* (Eds. I. Peretz & R.J. Zatorre) Science and Medical Publications. Oxford University Press, Oxford. 267–293

Blandin, Y., Lhuisset, L. & Proteau, L. (1999) Cognitive processes underlying observational learning of motor skills. *Quarterly Journal of Experimental Psychology: Human experimental psychology*, 52(A). 957–979

Blood, A., 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*. Vol. 2. No 4. 382–387

Blood, A. & 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*. Vol. 98. No 20. Sept. 25. 11818–11823

Bonanno, G.A., Papa, A., Lalande, K., Westphal, M. & Coifman, K. (2004) The importance of being flexible: the ability to both enhance and suppress emotional expression predicts long-term adjustment. *Psychological Science*. Vol. 15. No 7. 482–487

Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Understanding another's expectation from action: The role of peripheral sensation. *Nature Neuroscience*. Vol. 8. No 10. 1295–1297

Bossius, T. (2003) *The Future in the Rear Mirror. Black metal and transculture. Young people, music and religion in a post-modern world*. Daidalos, Gothenburg

Brattico, E. Näätänen, R. & Tervaniemi, M. (2002) Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception*. Winter. Vol. 19. No 2. 199–222

Bregman, A.S. (1990) *Auditory Scene Analysis – perceptual organization of sound*. MIT Press, Cambridge, Massachusetts

Brown, S., Martinez, M., Hodges, D., Fox, P. & Parsons L. (2004a) The song system of the human brain. *Cognitive Brain Research*. Vol. 20. No 3. 363–375

Brown, S., Martinez, M. & Parsons, L. (2004b) Passive music listening spontaneously engages limbic and paralimbic systems**.** *Neuroreport*. Vol. 15. No 13. September 15. 2033–2037

Bruner II, G.C. (1990) Music mood and marketing. *Journal of Marketing*. Vol. 54. No 4. 94–104

Bryngelsson, P. (2006) *Filmmusik – det komponerade miraklet*. Liber, Malmö

Bråten, S. (1999) *Intersubjective Communication and Emotion in Early Ontogeny*. Cambridge University Press

Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language*. Vol. 89. No. 2. 370–376

Burgess, N. (2002) The hippocampus, space, and viewpoints in episodic memory. *The Quarterly Journal of Experimental Psychololgy*. Vol. 55. No 4. 1057–1080

Burgess, N. (2006) Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Science*. Vol 10. No 12. 551–557

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A. & Rosen, B.R. (2002) Dorsal anterior cingulate cortex: a role in award-based decision making. *Proceedings of the National Academy of Sciences*, 99. 523–528

Bülthoff, H.H. & Christou, C.G. (2000) The perception of spatial layout in a virtual world biologically motivated computer vision: *First IEEE International Workshop*, BMCV Seoul, Korea, May, Proceedings, Vol. 811. Springer Berlin / Heidelberg. 10–19

Canetti, E. (1993) *Crowds and Power*. The Continuum Publishing Corporation, New York

Chafee, M.V. & Goldman-Rakic, P.S. (2000) Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. *Journal of Neurophysiology.* Vol. 83. No 3. 1550–66

Charles, M. (2005) Patterns: Basic unites of emotional memory. *Psychoanalytic Inquiry,* 25. 484–505

Chen, J., Penhune, V. & Zatorre, R.J. (2005) Tapping in synchrony to auditory rhythms. *Annals of the New York Academy of Sciences.* Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 400–403

Clegg, B.A., DiGirolamo, G.J. & Keele, S.W. (1998) Sequence learning. *Trends in Cognitive Science.* Vol. 2. No 8. 275–281

Clynes, D.M. (1977) *Sentics: The Touch of Emotion.* Anchor Press/Doubleday, New York

Cook, N. (1990) *Music, Imagination, and Culture.* Clarendon, Oxford

Cross, E.S., Hamilton, A.F.D.C., & Grafton, S.T. (2006). Building a motor simulation de novo: observation of dance by dancers, *NeuroImage.* Vol. 31. 1257–1267

Cross, I. (2003) Music, cognition, culture and evolution. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre) Oxford University Press, New York. 52–56

Csikszentmihályi, M. (1990) *Flow – the psychology of optimal experience.* Harper & Row Publishers, New York

Cumming, N. (2000) *The Sonic Self – musical subjectivity and signification.* Indiana University Press, Indianapolis; Bloomington

Cunningham, J.G. & Sterling, R.S. (1988) Developmental changes in the understanding of affective meaning in music. *Motivation and emotion.* Vol. 12. No.4. 399–413

Cutting, J.E. & Kozlowski, L.T. (1977) Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society.* 9. 353–356

Dallos, P. (1992) The active cochlea. *Journal of Neuroscience.* Vol. 12. No12. 4575–4585

DaMatta, R. (1991) *Carnivals, Rogues, and Heroes – an interpretation of the Brazilian dilemma.* University of Notre Dam Press, Indiana

Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y. & Iacoboni, M. (2006) Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience.* Vol. 9. No 1. January. 28–30

Darwin, C. (1872) *The Expression of the Emotions in Man and Animals.* J. Murray, London.

Davies, J.B. (1978) *The Psychology of Music.* Huthinson, London

Davies, M. & Troupe, Q. (1990) *Miles – the autobiography.* Pan Books, London.

Davies, S. (2001) Philosophical perspectives on music's expressiveness. In *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press. Oxford. 23–44

Dawkins, R. (2006) *The Selfish Gene.* Oxford University Press, Oxford. Chapter 11

De Gelder, B. & Hadjikhani, N. (2006) Non conscious recognition of body language. *Neuroreport.* Vol 17. No 6. 583–586

Deleuze, G. (1989) *The Time Image.* University of Minnesota Press, Minneapolis

Deutsch, D. (1999) *The psychology of Music.* Academic Press, San Diego

Dhamala, M., Pagnoni, G., Wiesenfeld, K., Zink, C., Martin, M. & Burns, G. (2003) Neural correlates of the complexity of rhythmic finger tapping. *NeuroImage.* Vol. 20. No 2. 918–926

Di Pellegrino, G., Fadiga L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992) Understanding motor events: a neurophysiological study. *Experimental Brain Research.* Vol. 91. No. 1. 176–180

Dimberg, U., Thunberg, M. & Elmehed, K. (2000) Unconscious facial reactions to emotional facial expressions. *American Psychological Society.* Vol.11. No 1. 86–89

Dissanayake, E. (1992) *Homo Aestheticus: where art comes from and why.* Free Press, New York

Dissanayake, E. (2000a) Antecedents of the temporal arts in early mother infant interaction. *The Origins of Music* (Eds. N. Wallin, B. Merker & S. Brown). The MIT Press, Cambridge, Massachusetts. 389–410

Dissanayake, E. (2000b) *Art and Intimacy: how the arts began.* University of Washington Press, Washington

Eco, U. (1979) *The Role of the Reader*. Indiana University Press, Bloomington

Ehrenfels, C. von (1988) *Foundations of Gestalt Theory.* (Ed. and translation B. Smith) Philosophia, Munich and Vienna. 82–117

Ekman, P. (1994) All emotions are basic. *The Nature of Emotions. Fundamental Questions* (Eds. P. Ekman & R.J. Davidson). New York University Press. New York. 15–19

Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T., Isham, E., Newman, E. & Fried, I. (2003) Cellular networks underlying human spatial navigation. *Nature.* Vol. 425. No 6954. 184–187

Fagius, J. (2001) *Hemisfärernas Musik*. Bo Ejeby förlag, Gothenburg

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. (2005) Parietal lobe: from action organization to intention understanding. *Science.* Vol. 308. No 5722. 662–667

Fraisse, P. (1982) Rhythm and Tempo. *The Psychology of Music.* (Ed. D. Deutsch) Academic Press, New York. 149–180

Freeman, W.J. (2000) A neurological role of music in social bonding. *The Origins of Music* (Eds. N.L. Wallin, B. Merker & S. Brown). MIT Press, Camebridge, MA. 411–424

Freud, S. (2004) *Mass Psychology and other Writings*. Penguin, London

Freyre, G. (1964) *The Masters and the Slaves*. Borzoi Books, New York

Fried, I. (2003) Cellular networks underlying human spatial navigation. *Nature.* Vol. 425. No 6954. 184–187

Fried, I., MacDonald, K.A. & Wilson, C.A. (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron*. Vol. 18. No 5. 753–765

Friedson, S. (1996) *Dancing Prophets: Musical experience in Tumbuka healing*. Chicago University Press, Chicago

From, F. (1971) *Perception of other people*. Colombia University Press, New York

Gabrielson, A. (2001) Emotions in strong experiences with music. *Music and Emotion – theory and research.* Eds. P.N. Juslin & J.A. Sloboda. Oxford University Press, Oxford. 431–449

Gabrielsson, A. & Lindstöm, S. (1993) On strong experiences in music. *Musik Psykologie.* 10. Florian Noetzel Verlag, Wilhelmshafen. 118–40

Gabrielsson, A. & Lindström, E. (2001) The influence of musical structure on emotional expression. *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press, Oxford. 223–248

Gabrielsson, A. & Lindström Wik, S. (2003) Strong experiences related to music: a descriptive system. *Musicae Scientiae.* Vol. 7. No 2. 157–217

Gallagher, S. & Jeannerod, M. (2002) From action to interaction. *Journal of Consciousness Studies.* Vol. 9. No 1. 3–26

Gallese, V. (2003a) The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology.* Vol. 36. No 4. 171–180

Gallese, V. (2003b) The manifold nature of interpersonal relations: the quest for a common mechanism. *Philosophical Transactions: Biological Sciences.* Vol. 358. No 1431. 517–528

Gans, D. (1991) Conversations with the dead. *The Grateful Dead Interview Book.* Citadel Underground, New York

Gazzaniga, M. (1985) *The Social Brain.* Basic Books, New York

Gelman, S. & Opfer J.E. (2002). Development of the Animate-Inanimate Distinction. *Blackwell Handbook of Cognitive Development* (Ed. U. Goswami). Blackwell Publishing. Oxford, UK.151–166

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences.* Vol. 7. No 7. 287–292

Gibson, E.J. (1953) Improvements in perceptual judgments as a function of controlled practice or training. *Psychological Bulletin.* No 50. 401–431

Gibson, J.J. (1986) *The Ecological Approach to Visual Perception.* Lawrence Erlbaum Associates, Hillsdale, New Jersey

Gnoli, R. (1956) *The Aesthetic Experience According to Abhinava Gupta.*: Chowkhamba Publications, Varanasi, India

Goldstone, R. (1998) Perceptual learning. *Annual Review of Psychology.* February, Vol. 49. 585–612

Granot, R. & Donchin, E. (2002) Do Re Mi Fa So La Ti – constraints, congruity, and musical training: an event related brain potentials study of musical expectancies. *Music Perception.* Summer. Vol.19. No 4. 487–528

Grèzes, J., Frith, C.D. & Passingham, R.E. (2004) Inferring false beliefs from the actions of oneself and others: an fMRI study, *Neuroimage.* Vol. 21. No 2. 744–750

Griffiths, T.D., Warren, J.D., Dean, J.L. & Howard D. (2004) "When the feeling's gone": a selective loss of musical emotion. *Journal of Neurology Neurosurgery and Psychiatry*, 75. 344–345

Gross, J.J. (2002) Emotional regulation – affective, cognitive and social consequences. *Physiopsychology,* 39. 281–291

Hall, D.A. (2003) Auditory Pathways: Are 'What' and 'Where' Appropriate? *Current Biology.* Vol. 13. No 10. 406–408

Halpern, A.R. (2001) Cerebral substrates of musical imagery. *The Biological Foundations of Music.* (Eds. R.J. Zatorre & I. Peretz) *The Annals of the New York Academy of Science.* Vol. 930. 179–192

Halpern, A.R. & Zatorre, R.J. (1999) When that tune runs through your head: a PET investigation of auditory imagery of familiar melodies. *Cerebral Cortex.* Vol. 9. No. 7 October. 697–704

Hanslick, E. (1986) *On the Musically Beautiful.* Hacket Publ., Indianapolis

Haslinger, B., Erhard, P., Altenmuller, E., Schroeder, U., Boecker, H. & Ceballos-Baumann, A.O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience.* Vol. 17. 282–293

Hatfield, E., Cacioppo, J.T. & Rapson, R. L. (1993) *Emotional Contagion.* Cambridge University Press, New York

Haueisen, J. & Knoesche, T.R. (2001) Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience.* Vol. 13. No 6. 786–792

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences.* Vol. 6. No. 6. 242–247

Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. (2003) Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience.* Vol. 15. No 5. 673–682

Hindemith, P. (1941) *The Craft of Musical Composition.* Associated Music Publishers, New York

Hjelmslev, L. (1953) *Prolegomena to a Theory of Language.* Baltimore: Indiana University Massachusetts

Huron, D. (2001a) Is music an evolutionary adaptation? *The Biological Foundations of Music,* Vol. 930. (Eds. I. Peretz & R.J. Zatorre) *Annals of the New York Academy of Sciences*. 43–61

Huron, D. (2001b) Tone and Voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception.* Vol. 19, No. 1. 1–64

Huron, D. (2006) *Sweet Anticipation: music and the psychology of expectation.* MIT Press, Cambridge, Massachusetts

Husserl, E. (1991) *On the Phenomenology of the Consciousness of Internal Time.* Kluwer, Dordrecht

Huxley, J. (1951) Knowledge, morality & destiny. *Psychiatry.* May 14. 129–151

Hyde, K.L., Zatorre, R.J., Griffiths, T.J., Lerch, J.P. & Peretz, I. (2006) Morphometry of the amusic brain: a two-site study**.** *Brain.* Vol. 129. No 10. 2562–2570

Iacobini, M. & Lenzi, L.L. (2002) Mirror neurons, the insula, and empathy. *Behavioural and Brain Sciences.* Vol. 25. 39–40

Iacobini, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C. & Rizzolatti, G. (2005) Grasping the intentions of others with ones own mirror neuron system. *PlosBiology.* March. Vol 3. No 3. 529–535

International Who is Who in Music (annual) Vol 2. Popular Music. (Ed. S. Tyler) Melrose Press Ltd., Cambridge, England

Ito, M. (1993) Movement and thought: Identical control mechanisms by the cerebellum. *Trends in Neurosciences.* Vol. 16. No 11. 448–450

Jacobs, B.L. & Fornal, C.A. (1997) Serotonin and motor activity. *Current Opinion in Neurobiology*. Vol. 7. No 6. 820–825

James, W. (1884a) On some omissions of introspective psychology. *Mind*. Vol. 9. 1–26

James, W. (1884b) What is emotion? *Mind*. Vol. 9. 188–205

Janata, P., Tillmann, B. & Bharucha, J.J. (2002) Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cognitive, Affective & Behavioral Neuroscience*. Vol. 2. No 2. 121–140

Janata, P. & Grafton, S.T. (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*. Vol. 6. 682–687

Janata, P., Birk, J.L., Tillmann, B. & Bharucha, J.J. (2003) Online detection of tonal pop-out in modulating contexts. *Music Perception*. Vol. 20. No 3. 283–305

Jeannerod, M. (1994) The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*. Vol. 17. No 2. 187–245

Jeannerod, M. & Frak, V. (1999) Mental imaging of motor activity in humans. *Current Opinion in Neurobiology*. Vol. 9. No 6. 735–739

Johannisson, K. (2001) *Nostalgia – en känslas historia*. Albert Bonnier, Stockholm

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*. 14. 201–211

Jones, M.R., Moynihan, N., MacKenzie, N. & Puente, J. (2002) Temporal aspects of stimulus driven attending in dynamic arrays. *Psychological Science*. Vol. 13. No 4. 313–319

Juslin, P.N. & Laukka, P. (2003) Communication of emotions in vocal expression and music performance: different channels, same code? *Psychological bulletin*. Vol. 129. No 5. 770–814

Juslin, P.N. & Laukka, P. (2004) Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*. Vol. 33. No 3. 217–238

Juslin, P.N. (2001) Communicating emotion in music performance: A Review and theoretical framework, in *Music and Emotion – theory and research* (Eds. P.N. Juslin & J.A. Sloboda) Oxford University Press, Oxford. 309–337

Juslin, P.N. & Sloboda, J.A. (2001) *Music and Emotion – theory and research.* Oxford University Press, Oxford

Jääskeläinen, I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levänen, S. Lin, F., May, P., Melcher, J. Stufflebeam, S., Tiitinen, H. & Belliveau, J.W. (2004) Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the New York Academy of Sciences.* Vol. 101. No 17. 6809–6814

Kant, I. (1952) *The Critique of Aesthetic Judgement.* Clarendon, Oxford

Kennerley, S.W., Sakai, K. & Rushworth, M.F. (2004) Organization of action sequences and the role of the pre-SMA. *Journal of Neurophysiology.* Vol. 91. No.2. 978–993

Kessler, E.J., Hansen, C., & Shepard, R.N. (1984) Tonal schemata in the perception of music in Bali and the West. *Music Perception*, 2. 131–165

Keysers C., Kohler E., Umiltà, M.A., Nanetti, L., Fogassi L. & Gallese V. (2003) Audiovisual mirror neurons and action recognition. *Experimental Brain Research.* Vol. 153. 628–636

Kilner, J.M., Vargas, C., Duval, S., Blakemore, S.J. & Sirigu, A. (2004), Motor activation prior to observation of a predicted movement. *Nature Neuroscience.* Vol. 7. 1299–1301

Kivy, P. (2002) *Introduction to a Philosophy of Music.* Clarendon, Oxford

Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology.* Vol. 55. No 3. 1027–1046

Knoblich, G. & Jordan J.S. (2003) Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology. Learning Memory and Cognition.* Vol. 29. No 5. 1006–1016

Koch, W.A. (1989) *The Nature of Culture.* Brockmeyer, Bochum.

Koelsch, S., Gunter, T.C., von Cramon, D.Y., Zysset, S., Lohmann, G. & Friederici, A. (2002) Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage.* Vol. 17. No 2. 956–966

Koelsch, S. & Mulder, J. (2002) Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology.* Vol. 113. No 6. 862–869

Koelsch, S. & Friederici, A.D. (2003) Toward the neural basis of processing structure in music. Comparative results of different neurophysiological investigation methods. *Annals of the New York Academy of Sciences.* Vol. 999. *The Neurosciences and Music.* 15–28

Koelsch, S., Grossmann, T., Gunter, T.C., Hahne, H., Schröger, E. & Friederici, A.D. (2003) Children processing music: electric brain responses reveal musical competence and gender differences. *Journal of Cognitive Neuroscience.* Vol. 15. No 5. 683–693

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

Koelsch, S., Fritz, T., v. Cramon, Y., Müller, K., & Friederici, A.D. (2006) Investigating emotion with music: An fMRI study. *Human Brain Mapping.* Vol. 27. No 3. 239–250

Koffka, K. (1935) *The Principles of Gestalt Psychology.* Lund Humphries, London

Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V. & Rizzolatti, G. (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science.* Vol. 297. No 5582. 846–848

Kosslyn, S.M. (1973) Scanning visual images: Some structural implications. *Perception and Psychophysics.* Vol. 14. No 1. 90–94

Kosslyn, S.M. (1975) Information representation in visual images. *Cognitive Psychology.* No 7. 341–370

Kreiman, G., Koch, C., & Fried, I. (2000). Imagery neurons in the human brain. *Nature.* Vol. 408. No 6810. 357–361

Kristeva, J. (1989) *Black Sun.* Colombia University Press, New York

Krumhansl, C.L. (1990) *Cognitive Foundation of Musical Pitch.* Oxford University Press, New York

Krumhansl, C.L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum.* Vol. 17. No 1. 53–80

Krumhansl, C.L. & Toiviainen, P. (2003) Tonal Cognition. *The Cognitive Neuroscience of Music.* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. 95–108

Körlin, D. (2005) *Creative Arts Therapies in Psychiatric Treatment – a clinical application of the Bonny Method of guided imagery and music (BMGIM) and creative arts groups.* Doctoral dissertation. Karolinska Institutet, Stockholm

Lahav, A., Boulanger, A., Schlaug, G., & Saltzman, E. (2005) The power of listening: auditory-motor interactions in musical training. *Annals of the New York Academy of Sciences.* Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 189–194

Lahav, A., Saltzman, E. & Schlaug, G. (2007) Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience.* Vol. 27. 308–314

Lakoff, G. & Johnson, M. (1980) *Metaphors We Live By.* University of Chicago Press, Chicago

Lakoff, G. & Johnson, M. (1999) *Philosophy in the Flesh. The embodied mind and its challenge to Western thought.* Basic Books, New York

Langer, S. (1942) *Philosophy in a New Key: a study in the symbolism of reason, rite, and art.* Harvard University Press, Cambridge, MA

Langer, S. (1953) *Feeling and Form.* Charles Scribner's Sons, New York

Large, E.W. & Kolen, J.F. (1994) Resonance and the perception of musical meter. *Connection Science.* Vol. 2. No. 6. 177–208

Laukka, P. (2004) *Vocal Expression of Emotion – discrete-emotions and dimensional accounts.* Acta Universitatis Upsaliensis, Uppsala

Lazarus, R. (1991) *Emotion and Adaptation*. Oxford University Press, New York

Le Bon, G. (1982) *The Crowd: a Study of the Popular Mind*. Atlanta: Cherokee Publishing Company

LeDoux, J. (1996) *The Emotional Brain* – the mysterious underpinnings of emotional life. Touchstone, New York

Lehrdal, F. (1988) Cognitive constrains on compositional systems. *Generative Processes in Music*. (Ed. John Sloboda). Oxford University Press. 231–259

Levine, M. (1995) *The Jazz Theory Book*. Sher Music. Peta Luma, CA. 183–192

Levison, J. (1996) *The Pleasures of Aesthetics*. Cornell University Press. Ithaca, NY

Lévi-Strauss, C. (1955) *Tristes Tropiques*. Plon, Paris

Lévi-Strauss, C. (1958) *Anthropogie Structurale*. Plon, Paris

Lévi-Strauss, C. (1968) *Mythologiques*. Plon, Paris

Lévi-Strauss, C. (1975) *The Raw and the Cooked. Introduction to a science of mythology*. Harper, New York

Levitin, D.J. (2005) Musical behavior in a neurogenetic developmental disorder – evidence from Williams Syndrome. *Annals of the New York Academy of Sciences* 1060. *The Neurosciences and Music II: From Perception to Performance*. (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) 325–334

Levitin, D.J., Menon, V., Schmitt, J.E., Eliez, S., White, C.D., Glover, G.H., Kadis, J., Korenberg, J.R., Bellugi, U. & Reiss A.L. (2003) Neural correlates of auditory perception in Williams Syndrome: An fMRI Study, *NeuroImage*. Vol. 18. No. 1. 74–82

Lhermitte, F., Pillon, B. & Serdan, M. (1986) Human autonomy and the frontal lobes: Part 1. Imitation and utilization behaviour: A neurological study of 75 patients. *Annals of Neurology*. Vol. 19. 326–34

Lipps, T. (1903) Einfühlung, innere Nachahmung und Organempfindung. *Archiv für gesamte Psychologie*. 1. 465–519

Lucas, G. (2002) *Os Sons do Rosario: O Congado Mineiro dos Arturos e Jatobá*. Editora UFMG, Belo Horizonte

Lucas, G. (2005). An ethnographic perspective of musical entrainment. *In Time with the Music: the concept of entrainment and its significance for ethnomusicology* (Eds. M. Clayton, R. Sager, G. Lucas & U. Will. *European Meetings in Ethnomusicology (special ESEM-Counterpoint volume).* Vol. 11. 96–99

Luu, P., & Pederson, S.M. (2004) The anterior cingulate cortex: regulating actions in context. *Cognitive neuroscience of attention* (Ed. M.I. Posner). Guilford, New York

Maidhof (2007) Neurophysiological correlates of error detection in performing musicians. Poster abstract at the Cognitive Neuroscience Society Annual Meeting Program 2007.

Malmgren, H. (2004) Why the past is sometimes perceived, and not only remembered. *Philosophical Communications, Web Series*

Manthey, S., Schubotz, R.I., & von Cramon, D. Y. (2003). Pre-motor cortex in observing erroneous action: An fMRI study. *Cognitive Brain Research.* Vol. 15. No 3. 296–307

Marcel, A.J. (1983) Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology.* Vol. 15. 238–300

Marshall, S. & Cohen, A.J. (1988) Effects of musical soundtracks on attitudes to geometric figures. *Music Perception,* 6. 95–112

Martin, H. (1996) *Charlie Parker and Thematic Improvisation.* Scarecrow Press, Metuchen N.Y.

Martin, J. (1978) *El Arte Flamenco de la Guitarra*, United Music Publishers Ltd, London.

Martinez, J.L. (1991) *Música & Semiotica: Um estudo sobre a questão da presentação na liguagem musical.* Potifícia Universidade Católica de São Paulo, São Paulo

Masaki, Y., Nishikawa, T., Ikejiri, Y., Tokunaga, H., Hatta, N., Uema, T., Kazui, H., Doronbekov, T.K., Ogino, A. Miyoshi, N., Tanii, H., Tanaka, T., Oku, N. & Takeda, M. (2006) Neural substrates of emotionally valenced episodic memory: A PET study using film stimuli. *Psychiatry and Clinical Neurosciences.* Vol. 60. No 1. 46–51

Mather, G. (2006) *Foundations of Perception*. Psychology Press, Hove

Matravers, D. (1998) *Art and Emotion*. Clarendon, Oxford

Mattheson, J. (1739) *Der volkommene Capellmeister*. Herold, Hamburg

Maturana, H. & Varela, F.J. (1987) *The Tree of Knowledge: The biological roots of human understanding*. Shambala, Boston

McAdams, S. & Matzkin, D. (2003) The roots of musical variations in perceptual similarity and invariance. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. 79–94

McAuley, J.D. (1995) *Perception of Time as Phase: Toward an adaptive-oscillator model of rhythmic pattern processing*. Technical Report No. 151 (Doctoral dissertation), Indiana University Cognitive Science Program

Meck, W. H. (1996) Neuropharmacology of timing and time perception. *Cognitive Brain Reseach*. Vol. 3. No 3–4. 227–242

Mega, M., Cummings, J. Salloway, S. & Malloy, P. (1997) The limbic system: an anatomic, phylogenetic, and clinical perspective. *The Neuropsychiatry of Limbic and Subcortical Disorders* (Eds. S. Salloway, P. Malloy & J. Cummings). American Psychiatric Press, Washington DC. 3–18

Meister, I.G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R. & Thron, A. (2004) Playing piano in the mind – an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research*. Vol. 19. No 3. 219–228

Meltzoff, A.N. & Moore, M.K. (1987) Imitation of facial and manual gestures by human neonates. *Science*. Vol. 198. 75–78

Meltzoff, A.N., & Moore, M.K. (1997). Explaining facial imitation: theoretical model. *Early Development and Parenting*. Vol. 6. 179–192

Menon, V. & Levitin, D.J. (2005) The rewards of music listening: Response an physiological connectivity of the mesolimbic system. *NeuroImage*. Vol. 28. No 1. 175–184

Merker, B. (2002) Music: The missing Humboldt system. *Musicae Scientiae*. Vol. 6. Spring issue. 3–21

Merleau-Ponty, M. (1996) *Phenomenology of Perception*. Routledge & Kegan Paul, London

Metzinger, T. & Gallese, V. (2003) The emergence of a shared action ontology: Building blocks for a theory. *Consciousness and Cognition*. Vol. 12. 549–571

Meyer, L.B. (1956) *Emotion and Meaning in Music*. University of Chicago Press, Chicago

Meyer, M., Steinhauer, K., Alter, K., Friederici, A.D. & von Cramon, D.Y. (2004) Brain activity varies with modulation of dynamic pitch variances in sentence melody. *Brain and Language*. Vol. 89. No. 2. 277–289

Molnar-Szakacs, I & Overy, K. (2006) Music and mirror neurons: From motion to e-motion. *Social Cognitive and Affective Neuroscience*. Vol. 1. No 3. 235–241

Monelle, R. (1992) *Linguistics and Semiotics in Music*. Harwood Academic Publishers, Edinburg

Muller-Preuss, P., Newman, J.D. & Jurgens, U. (1980) Anatomical and physiological evidence for a relationship between the “cingular" vocalization area and the auditory cortex in the squirrel monkey. *Brain Research*. Vol. 202. No 2. 307–315

Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J.G., Henning, J. & Seifritz, E. (2003) Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale, *Neuroimage*. Vol. 20. No 1. 429–434

Nattiez, J-J. (1975) *Fondements dune Semiologie de la Musique*. Union Générale d’Editions, Paris

Neisser, U. (1976) *Cognition and Reality*. Freeman, San Francisco

Newman-Norlund, R.D., van Shie, H.T., van Zuijlen, A. & Bekkering, H. (2007) The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*. Vol. 10. 817–818

Niedenthal, P.M. & Showers, C. (1998) The perdeption and processing of affective information and its influences on social judgement. In *Emotion and social Judgements* (Ed. J.P. Forgas) Pergamon Press, Oxford, England. 125–153

North, A.C. & Hargreaves, D.J. (1999) Music and driving game performance. *Scandinavian Journal of Psychology* Vol. 40. No 4. 285–292

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. (2001) "Primitive intelligence" in the auditory cortex. *Trends in Neurosciences,* Vol. 24. Issue 5. 283–288

Oatley, K. & Jenkins, J. (1996) *Understanding Emotions.* Blackwell, Cambridge, Massachusetts

Ochs, P. (2004) *Peirce, Pragmatism and the Logic of Scripture.* Cambridge University Press, Cambridge

Ogava, H., Ito, S. & Nomura, T. (1985) Two distinct projection areas from tongue nerves in the frontal operculum of macaque monkeys as revealed with evoked potential mapping. *Neuroscientific Research.* Vol. 2. No. 6. August. 447–59

Pallesen, K.J., Brattico, E., Bailey, C., Korvenoja, A., Koivisto, J., Gjedde, A. & Carlson, S. (2005) Emotion processing of major, minor and dissonant chords: an fMRI study. *The Neurosciences of Music. Annals of the New York Academy of Science.* Vol. 1060. 450–453

Panksepp, J. (1995) The emotional sources of chills induced by music. *Music Perception.* Vol. 13. No 2. 171–207

Panksepp, J. (1998) *Affective Neuroscience – the foundations of human and animal emotions.* Oxford University Press. New York

Panksepp, J. & Bernatzky, G. (2002) Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes.*Vol. 60. No. 2. 133–155

Pape, H. (2004) Peirce und seine nachfolger. *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlage von Natur und Kultur* (Eds. R. Posner, K. Robering & T.A. Sebeok). Teilband 2. De Gruyter, Berlin; New York. 2016–2040

Papousek, H. (1996) Musicality in infancy research: biological and cultural origins of early musicality. In *Musical Beginnings* (Eds. I. Delíege & M. Sloboda.) Oxford University Press, Oxford

Parsons, L.M. (2003) Exploring the functional neuroanatomy of music performance, perception and comprehension. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre) New York University Press. Chap. 17

Peirce, C.S. (1885) On the Algebra of Logic. A contribution to the philosophy of notation. *American Journal of Mathematics*. Vol. 7. No 2. January

Peirce, C.S. (1931–58) *Collected Papers*. (Eds. C. Hartshorn, P. Weiss & A.W. Burks). Harvard University Press. Cambridge, Massachusetts

Peirce, C.S. (1967) *Annotated Catalogue and the Papers of Charles Sanders Peirce* (Ed. R. S. Robin). Amhearst, MA

Peirce, C.S. (1980–2000) Harvard lectures on the logic of science. *The Writings of Charles Sanders Peirce*. Vol. 1. (Ed. M. Fisch)

Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I., Mottonen, R., Tarkiainen, A. & Sams, M. (2005) Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroimage*. Vol. 16. No 2. 125–128

Peretz, I. (2003) Brain specialization for music: New evidence from congenital amusia. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. 192–203

Peretz, I., Gagnon, L. & Bouchard, B. (1998) Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*. Vol. 68. No. 2. 111–141

Peretz, I., Brattico, E., & Tervaniemi, M. (2005) Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology*. Vol. 58. No 3. 478–82

Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E. & Evans, A. (1999) Localization of cerebral activity during simple singing. *NeuroReport*. Vol. 10. No 18. 3979–3984

Posner, M.I. & Dehaene, S. (1994) Attentional networks. *Trends in Neuroscience*. Vol.17. No 12. 75–79

Posner, R. (1989) What is culture? Toward a semiotic explanation of anthropological concepts. *The Nature of Culture* (Ed. W.A. Koch) Brockmeyer, Bochum. 240–295

Posner, R., Robering, K. & Sebeok, T.A. (2004) *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlage von Natur und Kultur.* De Gruyter, Berlin; New York.

Pratt, C. (1931) *Emotion and Meaning in Music: a study of psychological aesthetics.* McGraw-Hill Book Company Inc., New York and London

Preston, S.D. & de Waal, F.B.M. (2002) Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences.* Vol. 25. Issue 1. 1–72

Prinz, J.J. (2004) *Gut Reactions – a perceptual theory of emotion.* Oxford University Press, New York

Quartz, S. & Sejnowski, T. (2002) *Liars, Lovers, and Heroes – what the new brain science reveals about how we become what we are.* HarperCollins Publishers, New York

Quirk, G.J. & Milad, M.R. (2002) Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature.* Vol. 420. Nov. 7. 70–74

Raz, A., Fan, J. & Posner, M.I. (2005) Hypnotic suggestion reduces conflict in the human brain. *Proceedings of the National Academy of the United States of America.* Vol. 102. No 28. 9978–9983

Redish, D. (1999) *Beyond the Cognitive Map – from place cells to episodic memory.* The MIT Press, Cambridge, Massachusetts

Repérant, J., Miceli, D., Vesselkin, N.P. & Molotchnikoff, S. (1989) The centrifugal visual system of vertebrates: A century old research reviewed. *International Review of Cytology.* Vol. 118. 115–171

Repp, B.H. (2001) Effects of music perception and imagery on sensorimotor synchronization with complex timing patterns. *The Biological Foundations of Music.* (Eds. R.J. Zatorre & I. Peretz) *Annals of the New York Academy of Sciences.* Vol. 930, June. 409–411

Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G. & Heider, H. (2000) Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport.* Vol. 11. No. 9. 1997–2000

Rizzolatti, G. & Craighero, L. (2004) The mirror-neuron system. *Annual Review of Neuroscience*. Vol. 27. 169–192

Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (1999) Resonance behaviors and mirror neurons. *Archives of Italian Biology*. May. 137 (2–3). 85–100

Rizzolatti, G., Fogassi, L. & Gallese, V. (2001) Perspectives – neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews/Neuroscience*. Vol. 2. Sept. 667–668

Rizzo (2006) *Göteborgs-Posten*, October 14

Rizzuto, D., Mamelak, A., Sutherling, W., Fineman, I. & Andersen, R.A. (2005) Spatial selectivity in human ventrolateral prefrontal cortex. *Nature Neuroscience*. Vol. 8. No 4. 415–417

Rolls, E.T. (2005) *Emotion Explained*. Oxford University Press, Oxford

Rolls, E.T. (2006) Brain mechanisms of emotion and decision taking. *International Congress Series*. Vol. 1291. June. 3–13

Rousseau, J.J. (1986) Essay on the origin of languages. *On the Origin of Language: two essays*. (Eds. J.H. Moran & A. Gode) University of Chicago Press, Chicago. 5–74

Routledge, London. 259–271

Ruchsbaum, B., Olsen, R., Koch, P. & Berman, K. (2005) Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron*. Vol. 41. No 4. 687–697

Ryle, G. (1949) *The Concept of Mind*. The University of Chicago Press, Chicago

Sacks, O. (1990) *Awakenings*. Harper Perennial, New York

Sacks, O. (2007) The Abyss – music and amnesia. *The New Yorker*. Sept. 24.

Saffran, J.R. (2003) Mechanisms of musical memory in infancy. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York

Saito, Y., Ishii, K., Yagi, K., Tatsumi, I. & Mizusawa, H. (2006) Cerebral networks for spontaneous and synchronized singing and speaking. *Cognitive Neuroscience and Neuropsychology*. Vol.17. No 18. 1893–1897

Sakharov, D.S., Davydov, V.I. & Pavlygina, R.A. (2005) Intercentral relations of the human EEG during listening to music. *Human Physiology.* Vol. 31. No 4. 392–397

Sartre, J.-P. (1958) *Being and Nothingness.* Methuen, London

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D.I. & Kanwisher, N. (2004) A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia.* Vol. 42. No 11. 1435–1446

Scherer, K.R. (1986) The vocal affect expression: A review and a model for future research. *Psychological Bulletin.* Vol. 99. No 2. 143–165

Schmahmann, J.D., Anderson, C. M., Newton, N. & Ellis, R. (2001) The function of the Cerebellum. *Consciousness & Emotion*. Vol. 2. No 2. 273–309

Schubotz, R.I., Sakreida, K., Tittgemeyer, M. & von Cramon, D.Y. (2004) Motor areas beyond motor performance: deficits in serial prediction following ventrolateral pre-motor lesions. *Neuropsycology.* Vol. 18. No 4. 638–645

Schultz, W. Dayan, P. & Montague, P.R. (1997) A neural substrate of prediction and award. *Science.* Vol. 275. March14. 593–599

Schupp, H.T., Junghofer, M., Weike, A.I. & Hamm, A.O. (2003) Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *Neuroreport.* Vol. 14. No 8. 1107–1110

Schurmann, M., Raij, T., Fujiki, N. & Hari, R. (2002) Mind's ear in a musician: where and when in the Brain. *NeuroImage.* Vol. 16. No 2. 434–440

Schütz, A. (1964) *Collected papers: studies in social theory.* Vol 2. (Ed. A. Brodersen). Martinus Nijhoff, Hague

Schönrich, G. (1990) *Zeichenhandeln: Untersuchungen zum Begriff einer semiotichen Vernunft im Ausgang von Charles Sanders Peirce.* Suhrkamp, Frankfurt am Main.

Scruton, R. (1983) Understanding music, *Ratio.* Vol. 25, No. 2. 97–120

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

Sebanz, N., Bekkering, H. & Knoblich, G. (2006) Joint action: bodies and minds moving together. *Trends in Cognitive Science.* Vol. 10 No. 2. 70–76

Shannon, C.E. & Weaver, W. (1949) *A Mathematical Model of Communication*. University of Illinois Press, Urbana, IL

Sheppard, R.N. & Metzler, J. (1971) Mental rotation of three-dimensional objects. *Science*. Vol. 171, No. 972. 701–703

Shin, J.C. & Ivry, R.B. (2002) Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition*. Vol. 28. No 3. 445–457

Simons, D.J. & Wang, R.F. (1998) Perceiving real-world viewpoint changes. *Psychological Science*. Vol. 9. No 4. 315–320

Smith, A.D. (2002) *The Problem of Perception*. Harvard University Press, Cambridge, Massachusetts

Smith, G.J. (1999) Teaching a long sequence of behavior using whole task training, forward chaining, and backward chaining. *Perceptual and Motor Skills*. Vol. 89. No 3. 951–965

Snyder, B. (2000) *Music and Memory*. MIT Press, Cambridge Massachusetts

Sonneby-Borgström, M. (2002) Automatic mimicry reactions as related to differences in emotional empathy. *Scandi–navian Journal of Psychology*. Vol. 43. No 5. 433–443

Spence, C. & Read, L. (2003) Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *American Psychological Society*. Vol. 14. No 3. 251–256

Spencer, H. (1857) The origin and function of music. *Fraser's Magazine*, 56. 367–408

St. Clair, R.N. & Bush, J.A. (2001) Transmission of values: the information age crisis in socialization. *Intercultural Communication Studies XI*. 67–69

Stefani, G. (1976). *Introduzione alla Semiotica della Musica*. Sellerio, Palermo

Stern, D. (1985) *The Interpersonal World of the Infant*. Basic Books, New York,

Stern, D. (1993) The role of feelings for an interpersonal self. *The Perceived Self* (Ed. U. Neisser). Cambridge University Press

Stern, D. (1995) *The Motherhood Constellation*. Basic Books, New York

Stern, D. (2004) *The Present Moment*. W.W. Norton & Company, New York, London.

Stravinky, I. (1936) *Cronicle of my Life*. Victor Gollancz Ltd., London

Sundqvist, F. (2003) *Perceptual Dynamics – theoretical foundations and philosophical implications of Gestalt psychology*. Acta Universitatis Gothoburgensis, Gothenburg

Tagg, P. (1979) *50 Seconds of Television Music – towards an analysis of affect in popular music*. Studies from the Department of Musicology. No 2. Gothenburg

Tarasti, E. (1994) *A Theory of Musical Semiotics*. Indiana University Press, Indianapolis; Bloomington

Tervaniemi, M. (2003) Musical sound processing: EEG and MEG evidence. *The Cognitive Neuroscience of Music*. (Eds. I. Peretz & R.J. Zatorre) Oxford University Press, New York. 294–309

Thomas, N. (1999) Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*. Vol. 23. No 2. 207–245

Tillmann, B. (2005) Implicit investigations of tonal knowledge in non-musician listeners. *Annals of the New York Academy of Sciences*. Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo). 100–110

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

Tillmann, B., Bharucha, J.J. & Bigand, E. (2003b) Learning and perceiving musical structures: Further insights from artificial networks. *The Cognitive Neuroscience of Music* (Eds. I. Peretz & R.J. Zatorre). Oxford University Press, New York. Chapter 8.

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D. & von Cramon, D.Y. (2006) Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage*. Vol. 31. No 4. 1771–1782

Tomkins, S. (1991) *Affect, Imagery and Consciousness*. Springer Publishing Company, New York

Towner, R. (1985) *Improvisation and Performance Techniques for Classical and Acoustic Guitar.* 21st Century Music Productions. Wayne, New Jersey

Trevarthen, C. (1992) An infant's motives for speaking and thinking in the culture. *The Dialogical Alternative.* (Ed. A.H. Wold.) Scandinavian University Press/Oxford University Press, Oslo and Oxford. 99–137

Trevarthen, C. (1999) Musicality and intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae.* Special issue. 155–215

Trivedi, S. (2003) The funerary sadness of Mahler´s music. *Imagination, Philosophy and the Arts. (Eds. M. Kieran & D. McIver Lopes)* Routledge, London. 259–271Tupinambá de Ulhôa, M. (1999) Métrica Derramada: Prosódia musical na canção brasileira popular. *Brasiliana – Revista da Academia Brasileira de Música.* Vol. 2. 48–56

Ullén, F., Bengtsson, S., Ehrsson, H.H. & Forsberg, H. (2005) Neural control of rhythmic sequences. *Annals of the New York Academy of Sciences*. Vol. 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo) Vol. 1060. 368–376

Ullsperger, M. & von Cramon, D. Y. (2003) Error monitoring using external feedback: Specific roles of the habenular complex, the reward System, and the cingulate motor area revealed by functional magnetic resonance imaging. *The Journal of Neuroscience.* Vol. 23 No 10. 4308–4314

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron.* Vol. 32. 91–101

Walton, K.L. (1988) What is abstract about the art of music? *Journal of Aesthetics and Art Criticism.* Vol. 46. No 3. 351–64

van Schie, H.T., Mars, R.B., Coles, M.G.H. & Bekkering, H. (2004), Modulation of activity in medial frontal and motor cortices during error observation, *Nature Neuroscience.* Vol. 7. No 5. 549–554

Wang, R.F. & Simons, D.J. (1999). Active and passive scene recognition across views. *Cognition,* Vol. 70. No 2. 191–210

Warren R.M. (1970) Perceptual restoration of missing speech sounds. *Science, 23*, Vol. 167. No. 3917. 392–393

Weinberger, N.M. (2004) Music and the brain. *Scientific American.* Nov. 89–95

Wessely, O. (1972) *Musik.* Habel, Darmstadt

Whitehouse, H. (2000) *Arguments and Icons: divergent modes of religiosity.* Oxford University Press, Oxford

Wicke, P. (2003) Popmusic in der Analyse. *Acta Musicologica.* Vol. 75. No 1. 107–126

Vickhoff, B. & Malmgren, H. (2005) Why does music move us? *Philosophical Communications.* Web Series, No. 34. Dept. of Philosophy, Göteborg University, Sweden. ISSN 1652–0459

Williams, L., Brown, K., Palmer, D., Liddell, B., Kemp, A., Olivieri, G., Peduto, A. & Gordon, E. (2006) The mellow years? Neural basis of improving emotional stability over age. *The Journal of Neuroscience.* Vol. 26. No 24. 6422–6430

Wilson, M. and Knoblich, G. (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletine.* Vol. 131. No. 3. 460–473

Wolfensteller, U., Schubotz, R.I., & von Cramon, D.Y. (2004) "What" becoming "Where": Functional magnetic resonance imaging evidence for pragmatic relevance driving pre-motor cortex. The Journal of Neuroscience, Nov. 17. Vol. 24. No 46. 10431–10439

Wraga, M., Creem, S.H. & Proffitt, D.R. (2000) Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology. Learning, Memory and Cognition.* Vol. 26. No 1. 151–68

Wright, B.A. & Fitzgerald, M. B. (2003) Sound – discrimination learning and auditory displays**.** *Proceedings of the 2003 International Conference on Auditory Display,* Boston, MA, 6–9 July. 228–232

Vuilleumier, P. (2005) How brains beware: the neural mechanism of emotional attention. *Trends in Cognitive Sciences.* Vol. 9. No 12. 586–595

Västfjäll, D. (2002) A review of the musical mood induction procedure. *Musicae Scientiae.* Special Issue. 173–211

Yabe, H., Matsuoka, T., Sato, Y., Hiruma, T., Sutoh, T., Koyama, S., Gunji, A., Kakigi, R. & Kaneko, S. (2005) Time may be compressed in sound representation as replicated in sensory memory. *Neuroreport.* Vol. 16 No 2. 95–98

Yumoto, M., Matsuda, M., Iteoh, K., Uno, A., Karino, S. Saitoh, O., Kaneko, Y., Yatomi, Y. & Kaga, K. (2005) Auditory imagery mismatch negativity elicited in musicians. *Neuroreport.* Vol. 16. No 11. 1175–1178

Zarate, J.M. & Zatorre R.J. (2005) Neural substrates governing audiovocal integration for vocal pitch regulation in singing. *Annals of the New York Academy of Sciences.* 1060. *The Neurosciences and Music II: From Perception to Performance* (Eds. G. Avanzini, S. Koelsch, L. Lopez & M. Manjo). 404–408

Zatorre, R.J. (2003) Neuronal specialization for tonal processing. In *The Cognitive Neuroscience of Music.* (Eds. I. Peretz & R.J. Zatorre) Oxford University Press, New York. 231–236

Zatorre, R.J. (2005) Music, the food of neuroscience? *Nature.* Vol. 434. No 7031. 312–316

Zatorre, R.J., Chen, J.L. & Penhune, V.B. (2007) When the brain plays music: auditory– motor interactions in music perception and production. *Nature Reviews. Neuroscience.* Vol. 8. No 8. 547–558

Zbikowski, L. (1998) Metaphor and music theory: reflections from cognitive science. *The Online Journal of the Society of Music Theory.* Vol. 4. No 1