The evolutionary origins of music

by

Sarah Wurz

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

at

Stellenbosch University

Department of Music

Supervisor: Prof Winfried Lüdemann

March 2009

`

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 16 February 2009

`

ABSTRACT

The evolutionary origins of music, defined as "an intentional action in which complex, learned vocalizations (and/or instrumentally produced sound) are combined with the movement of the body in synchrony to a beat" is investigated through an appraisal of the musilanguage theory and relevant literature.   The biological adaptations allowing the production and perception of music are identified and their evolutionary histories investigated. The critical adaptations that made rhythmical body movement possible evolved around 1.6 million years ago. These include habitual bipedalism and changes in the vestibular system. There is almost no fossil evidence to inform on the timing and nature of the complex, learned vocalization. However, that the thoracic vertebrate canal had modern proportions by 600 000 years ago indicates that archaic humans were able to achieve the respiratory control necessary to sing. The size of this canal is a proxy for the number of nerve cells that control respiration via the intercostal and abdominal muscles. Musicality is essential to the human mind. Infants are born with rudimentary musical skills with regard to melody, temporal sequences and vocal and bodily imitation. These capabilities are central to the newborns' innate ability to elicit care by synchronizing their vocal and bodily actions with that of the caregivers. Musical rhythm is further used to entrain bodily and neural oscillations and this permit the creation of trust and social bonding. It is concluded that protomusic developed between 1.6 million and 600 000 years ago. Protomusic consisted of entrained rhythmical whole body movements initially combined with grunt-like vocalizations. The evidence investigated cannot be used to infer the origins of modern music.

KEYWORDS: Music, Evolution, Synchronisation, Melody, Dance, Bipedality, Vestibular system, Thoracic vertebrate canal, Infant-directed communication, Neural entrainment

`

OPSOMMING

Die evolusionêre oorsprong van musiek, hier gedefinieer as "'n intensionele aksie waarin komplekse, aangeleerde vokalisasie (en/of instrumenteel geproduseerde klank) gekombineer word met die sinkroniese beweging van die liggaam tot 'n musikale pols" word ondersoek deur die 'musitaal' teorie en ander relevante literatuur te bestudeer. Die biologiese aanpassings wat die produksie en persepsie van musiek moontlik maak word geïdentifiseer en hulle evolusionêre agtergrond word ondersoek. Die kritiese aanpassings wat ritmiese liggaamsbewegings toelaat het ongeveer 1.6miljoen jaar gelede evolueer. Hierdie aanpassings sluit volkome bipedaliteit en veranderinge in die vestibulêre sisteem in. Daar bestaan feitlik geen fossiel bewyse waarvan die ontstaan en aard van komplekse, aangeleerde vokalisasie afgelei kan word nie. Desnieteenstaande, die torakale werwelkanaal het teen 600 000 jaar gelede moderne proposies ontwikkel en dit dui daarop dat argaïese mense die nodige respiratoriese kontrole kon uitoefen om te kan sing. Die grootte van hierdie kanaal is 'n aanduiding van die aantal senuselle wat respirasie beheer deur middel van die interkostale en abdominale spiere. Musikaliteit is essensieel tot menslike denke. Babas word gebore met basiese musikale vermoëns sover dit melodie, temporale sekwense en vokale en liggaamlike nabootsing aan betref. Hierdie vermoëns staan sentraal in pasgeborenes se ingebore aanleg om versorging te ontlok deur hul vokale en liggaams bewegings te sinkroniseer met dié van hul versorgers. Musikale ritme word verder gebruik om liggaams en neurale ossilasies met mekaar te sinkroniseer. Dit lei tot die vorming van vertroue en sosiale gebondenheid. Daar word tot die gevolgtrekking gekom dat protomusiek tussen 1.6 miljoen en 600 000 jaar gelede ontwikkel het. Protomusiek het bestaan uit doelbewus gesinkroniseerde ritmiese beweging van die liggaam wat aanvanklik met "steun"-agtige vokalisasies gepaard gegaan het. Die bewyse was nagevors is werp nie lig op die ontstaan van moderne musiek nie.

SLEUTELWOORDE: Musiek, Evolusie, Sinkronisasie, Melodie, Dans, Bipedaliteit, Vestibulere sisteem, Torakale werwelkanaal, Babagerigte kommunikasie, Neurale sinkronisasie

`

ACKNOWLEDGEMENTS

`

THE EVOLUTIONARY ORIGINS OF MUSIC

TABLE OF CONTENTS

`

`

`

CHAPTER 1: INTRODUCTION

1.1 THE AIM OF THE STUDY

For more than a century it has been proposed that modern music may have its origins in a musical protolanguage (Darwin 1871). Two publications, a book by Mithen (2005) and a thesis by Morley (2003), recently undertook in-depth investigation of Darwin's idea that music shares its roots with language in what has been termed 'musilanguage' (Brown 2000). The purpose of this study is to investigate the origins of music. The musilanguage theory as formulated by Mithen and Morley is used as starting point to identify the components suitable to a biomusicological investigation of the origins of music. These components are further developed through a literature review that takes the perspectives from various disciplines, including musicology, palaeoanthropology, brain science, and the developmental and biological aspects of psychology into account. This information is used to infer the nature and timing of protomusic and modern music

In this Chapter the research problem (1.2.) is introduced by discussing the theoretical context in which a common origin for language and music has been proposed (1.2.1). The main points from the musilanguage theory that are selected for research into the origins of music are given in Section 1.2.2. The examination of the origins of music is organised around four research questions originating from the musilanguage theory and these questions are formulated in 1.2.3. In Section 1.3 the research hypothesis is given while Section 1.4 presents an outline of the Chapters.

1.2 RESEARCH PROBLEM

1.2.1 The origins of music from a musilanguage point of view

The evolutionary origins of music has received intensive scientific investigation only for the past ten years and then virtually exclusively in the context of the evolution of language. Since the 1600's intellectuals like Descartes, Rousseau, Spencer and Darwin (Besson & Friederici 2005:57) considered the origins of music and language to be linked. Some hypothesise that language preceded music in evolution (Calvin 1996; Patel 2006), while others argue that music was the precursor of language (Vaneechoutee and Skoyles 1998; Benzon 2001). The most commonly proposed theory sees a common evolutionary foundation for language and music (Scherer 1991; Wilkens & Wakefield, 1996: Brown 2000, 2006; Morley 2002, 2003; Mithen 2005). This is known

`

as the 'musilanguage' theory (Brown 2000, 2006). The relationship between language and music is an extremely broad research topic that encompasses interdisciplinary literature that links studies in musicology, acoustics, linguistics, literary studies, philosophy, psychology and anthropology (Feld & Fox 1994: 26). This means that systematic multidisciplinary investigation of the topic is a mammoth task.

Mithen's "The Singing Neanderthals: The Origins of Music, Language, Mind and Body" is the first "book-length exposition of Darwin's musical protolanguage" (Fitch 2005b:288). Mithen proposes that tonal affective primate calls evolved into fully-fledged language and music between 1.6 million and 50 000 years ago. In this tonal, rhythmic, bodily entrained "protolanguage" the communication of emotion was crucial. Mithen prefers not to use the terms protolanguage or musilanguage to describe this communicative system because the use of 'language' in these terms is misleading (2005:26). The onomatopoeic term 'Hmmmmm' is used to refer to the single precursor of music and language. The precursor is termed Hmmmmm because it was <u>h</u>olistic, <u>m</u>anipulative, <u>m</u>ulti-<u>m</u>odal, <u>m</u>usical and <u>m</u>imetic.

A somewhat different trajectory for the origins of music and language is suggested by Morley (2003) who also proposes that a musical communication system existed, but suggests that it evolved between 1.75 mya and 300 000 years ago (Morley 2002:208). He comes to this conclusion from a multidisciplinary PhD study on the origins of music. His focus is on the capabilities that allowed human ancestors to produce complex vocalizations and rhythmic movements. He discusses broadly the same elements as Mithen and suggests that melody, rhythm, tonal vocalisation and the ability to gain affective information from these musical aspects were the key elements of a musical protolanguage (Morley 2003:217). The evolution of abilities that allowed the communication of corporeal and rhythmic expressions of affect would have been interdependent with the evolution of prosodically-contoured vocalisations. The musical protolanguage described by Morley and Mithen has rhythmic, melodious vocal and corporeal expression in common with music, and prosodic expression in common with modern language.

Prosody, expressed through vocalisation, is where music and language converge in the musilanguage theory. Brown (2000) contends that the intonational concerns for melody, rhythm, and phrasing in speech strongly parallel those in music. Mithen follows this idea by stating that: "…the melodic and rhythmical nature of spoken utterances; when the prosody is intense, speech sounds highly musical" (Mithen 2005:24). Morley (2003: 137) also accepts this proposition by

`

highlighting the elements of melody, tempo and pitch as common elements in music and prosody. Speech prosody occurs because the emotional state of a person (or animal) influences the acoustic qualities of a voice. Prosodic indicators include completely non-linguistic verbal utterances such as crying or laughing or paralinguistic indicators such as pitch and tone, rhythm, intonation patterns, stress, timing and differential pausing in speech that indicate emotional and attitudinal states (Monnot *et al* 2004:519). There is a difference between prosody for linguistic functions, for example to recognize the difference between a statement and a question, and non-linguistic prosody.

The definition of music is discussed in Chapter 2, but it should be noted here that music involves a much more extensive usage of pitch and rhythm than prosody. For example, musical melody consists of successive pitch changes and intervals while prosody uses high-level, rising and low dipping or falling contours. In prosody a descending pitch often marks the end of phrases in sentences (Thompson *et al* 2004:38). Despite extensive research, the melody aspect of speech is still poorly understood (Xu 2005:221). Mithen (2005: 24) and Morley (2003:218) both include 'rhythm' when describing the aspects that music and prosody have in common. However, it is most probable that the rhythmical properties of music and language are not the same (Bispham 2006b:127) as the 'beats' of speech do not form a regular pulse (Patel 2006:100). Music and prosodic speech are both "systems of expressively intoned sound" (Brown 2000:279), but both are complex phenomena. This necessitates in depth study of both phenomena. In this study the choice has been made to focus only on the musical aspects of musilanguage to investigate the origins of music. To adequately investigate how prosody relates to musical expression would have required extensive examination of a large body of literature. Such a broad comprehensive investigation was beyond the scope of this thesis.

In this study music is regarded as a complex phenomenon that is adaptive. Steven Pinker famously asserted in "How the mind works" (1997) that music is non-adaptive because there is no relationship between music and biological competencies such as accurate perception, vision, social reasoning. However, it is premature to argue that music's origin is unknowable and that it is not adaptive (McDermott 2008:287). The strongest evidence against Pinker's "cheesecake" hypothesis (Fitch 2006a:199) is that music, like language, is present in and fundamental to all known human societies (Cross 2003:2). The understanding that music is an embodied, as well as a cognitive activity (Gamble 2006:109) is crucial in the search for its evolutionary roots and therefore the

`

origins of music are investigated from a biomusicological point of view. The field of biomusicology or "biology and the evolution of music" (Fitch 2006a:174) has recently received a surge of interest (Fitch 2006a:174). A variety of capabilities and mechanisms are involved in perceiving and producing music and each of these may have a different evolutionary history. Therefore, as Fitch (2006a:174) points out, it does not make much sense to simply put "When did music evolve?" or "What is music for?" as research questions as the answer to these questions are not unitary. The approach taken in this study is to investigate the origins of music through a number of musical capabilities and their biological correlates.

There are practically no extended studies that investigate the evolutionary roots of music *per se*. The multidisciplinary extended analyses of Mithen and Morley's therefore presents an ideal opportunity to study the roots of music. Their publications discuss the evolution of musilanguage from somewhat different perspectives and this allows the investigation of the origins of music from a broad multidisciplinary basis. The research method followed here involves a multidisciplinary literature review. The musilanguage theory is examined with the aim of isolating and further investigating those biological adaptations that relate to the perception and production of music. The evolution of grammatical language is not addressed whatsoever. How the origins of music may relate to grammatical language is touched upon in a very cursory manner within the context of analysing musilanguage.

The following propositions, put forward by Mithen and Morley respectively in their discussions of the evolution of musilanguage, are selected to investigate the origins of music.

1.2.2 The aspects of musilanguage relevant to the evolutionary origins of music

1. Mithen and Morley tie the ability to vocalize musically to the descent of the larynx, to changes in the hypoglossal canal and to the flexion of the basicranium. The size of the thoracic canal can be related to breathing capabilities and this is also related to the evolution of musical expression by both authors. They further discuss adaptations related to bipedalism in the context of rhythmical movement. These biological parameters and their relevance for the origins of music will be critically evaluated, and the significance of other indicators, for example the vestibular system of the inner ear will be discussed.

2. Morley discusses the evolution of the neural substrates of musilanguage in detail. He concludes that between 1.75 million years ago and 300 000 years ago increased vocal

control was accompanied by lateralization of prosody and emotional functions in the right hemisphere. The neurological pathways to control laryngeal and orofacial muscles are related to left hemisphere mechanism. These left hemisphere developments are tied to motor sequences associated with rhythmic behaviour. Mithen's discussion of neural relationships underlying musilanguage is not as detailed as that of Morley, but his view of musical protolanguage as being underpinned by relatively discrete modules in the brain is dominant throughout. His modular view of brain evolution guides him to propose that Neanderthals possessed superior musical skills, but inferior linguistic capabilities. The evidence for lateralization, modularity and its relationship to the evolution of musical capabilities will be investigated.

3. The sing-song way in which caregivers communicate with infants, termed Infant-directed communication, is potentially informative on the origins of music. Infants naturally perceive and entrain with this melodic, slow and repetitive communication and this is why it has been suggested that the capacity to perceive melody and rhythm and to gain affective information from it is innate in newborns. Both Mithen and Morley discuss its possible relevance to the evolution of musilanguage. Mithen provides an interesting chapter on Infant-directed communication and he investigates whether it contains mechanisms that once belonged to a musical ability that was used to regulate social relationships and emotional states. Morley discusses this aspect less comprehensively, but he provides a clear hypothesis for further investigation – that the infant's sense of rhythm and melody may have a basis in hereditary factors. The musical capabilities of pre-natal and very young infants are researched in terms of melody, rhythm and motivational components.

4. It is ofen argued that musicality evolved as a means to express and induce emotions and that the "emotional" nature of both music and prosody indicate that they share a common basis in musilanguage. Mithen explicitly links the evolutionary significance of music and prosody to the expression and generation of emotion and Morley, along the same lines, sees the emotive elements of vocalisation in music and prosody as a fundamental component of evolutionary fitness. However, it is problematical to study "emotion" because so many subjective meanings can be ascribed to it. The approach followed in this study is to isolate those elements of "emotion" amenable to biological investigation.

`

1.2.3 Formulation of the research questions

This study will be organised around four research questions:

1. Which biological aspects leave traces in the fossil record that can be reliably related to music production and perception, and what implications do these finds have for the origins of music?
2. Is there evidence from the brain sciences that can inform on the evolutionary origins of music?
3. Are infants born with innate musical capabilities and is Infant-directed communication relevant to the study of the origins of music?
4. Are there certain "emotional" responses to music that are evolutionary meaningful?

1.3 FORMULATION OF THE HYPOTHESIS

This study is aimed at identifying the biological components related to music making and perception and to interpret their evolutionary histories in terms of the origins of music. This research is based on the hypothesis that the propositions made by Mithen and Morley about the evolutionary origins of music are open to criticism and that they can be revised in the light of relevant literature. The propositions in question are that there is sufficient evidence from the fossil record to reliably infer the evolutionary origins of the production and perception of music; that the neuro-scientific literature supports the notion that the evolution of lateralisation and modularisation of the brain underlies musical capabilities; that infants have innate musical capabilities; and that there are certain "emotional" responses to music that are meaningful from an evolutionary perspective. These propositions indicate that the capabilities for human music making, in the form of entrained movement and complex learned vocalisation, evolved between 1.6 million years ago and 50 000 years ago.

1.4 CHAPTER OUTLINE

Chapter 2 introduces the concepts of the analytical framework that is used to investigate the origins of music. Music and the musical elements of pitch and rhythm are defined. The biological apparatus that relate to the production and perception of musical sound and the execution of entrained

`

movements are also described. In Chapter 3 the musilanguage theories of Mithen and Morley are discussed in terms of their primate origins, fossil evidence and rationales. One of the objectives of this chapter is to discuss the fossil evidence for the biological elements underlying musical expression identified in Chapter 2. From the analysis in this Chapter three critical components that need to be further investigate are defined - these include brain evolution, the innate capabilities of pre-natal and very young infants and the "emotional" response to music.

Chapter 4 begins by analysing Morley's and Mithen's ideas on how brain evolution relates to musilanguage. Morley's path for the evolution of the brain is based on the current neural processing of prosodic, tonal, rhythmic and emotional aspects of music and language. He proposes that the most ancient substrate for musilanguage relates to complex vocal emotional expression and comprehension. In Mithen's theory, specific modules related to pitch, rhythm and emotion enable Hmmmmm and it is only when these previously isolated brain modules become integrated that modern music (and syntactical language) developed. These propositions are investigated in the context of complementary neuropsychological and brain imaging studies on pitch, rhythm and emotion. The elements pertinent to the evolution of musical capabilities are extracted from these studies.

The musical behaviours of infants are given much significance in the theories of Mithen and Morley. For example, the "innate" preference of infants for infant-directed communication and their ability to comprehend musical emotional utterances are regarded behavioural fossils of a musical communication system in our ancestors' past. The relevance of infant musical behaviour for the evolutionary origins of music is examined and further developed in Chapter 5.

Mithen's characterisation of Pleistocene hominins as highly emotional and therefore highly musical (Mithen 2006:100) may be somewhat extreme, but he, like Morley weaves emotion through every aspect of musilanguage. The premise is that music is *for* emotional expression. In Chapter 6 it is argued that the most advantageous way to approach "emotion" in a study of musical origins is to regard it from a biological point of view. For this reason the biopsychological effect of music, including the physiological and neurochemical impact of music, is investigated.

This study concludes with Chapter 7, in which the research questions will be discussed and the research hypothesis rejected or accepted according to the information presented.

`

CHAPTER 2: THE ELEMENTS OF MUSIC AND ITS BIOLOGICAL CORRELATES

2.1 INTRODUCTION

In a study of the evolutionary origins of music, adequate definition, as Bispham (2006a: 589) and Morley (2006:101) emphasise, is essential. For this reason the definition of "music" is discussed in Section 2.2. The elements of music that are widely discussed in the context of the origins of music - pitch and rhythm are described in section 2.3. The research questions set in Chapter 1 revolve around identifying, describing and investigating the evolution of the capabilities underlying musical ability. The biological components that make it possible to perceive and express musically involve adaptations that enable vocalisation, respiration, hearing and movement. The aim in Section 2.4 is to describe these adaptations and to identify which of these are conserved in the fossil record. This discussion of the biological basis of music provides a concrete foundation for the investigation of the evolutionary origins of music in the following Chapters.

2.2 MUSIC

Definitions of music are hard to come by. As Fitch (2006a:181) mentions: "…no uncontroversial definition [of music] is currently accepted…". Huron's (2001:44) comment that the "nebulous rubric music" may represent several adaptations, which involved complex co-evolutionary patterns with culture, rings true. At its most basic, music is organised sound (Varèse in Levitin 2006:14). Human music "relies on a discretization of both pitch and time" (Fitch 2006a:179) and thus involves the components melody and rhythm. Human music is also be distinguished by unique timbre (Krumhansl & Iverson 1992), but neither Mithen nor Morley elaborate on its evolutionary significance. It is an understudied component of music evolution that needs much deeper investigation.

Music involves song and instrumental music and Fitch (2006a:195) suggests that song and bimanual drumming are probably the oldest forms of music. An "objective" (Fitch 2006a:182) definition of song that can be investigated from a comparative biological point of view, is

`

"complex, learned vocalization". As will be discussed in Chapter 3, animals like birds and whales also produce melodies that can be described as "complex, learned vocalization" in that they are newly constructed, manipulated and learnt, but this does not mean that human song and animal song have the same evolutionary root (or are homologous). Animal song and human song have different evolutionary histories and are therefore evolutionary analogs and convergent adaptations (Fitch 2006a:182, but see Hauser & McDermott 2003:667). Instrumental music can be defined as "….the use of the limbs or other body parts to produce structured, communicative sound, possibly using additional objects…" (Fitch 2006a:183). The rhythmical aspect of music involves the production of a regular beat to which movements can be adapted. Human music is unique in the sense that it contains a regular beat to which bodily movements can be synchronised (Patel 2006:100) and entrained (Bispham 2006a: 589).

From an evolutionary point of view music in its broadest conception and expression is relevant. Both Mithen (2005:11-12) and Morley (2003:4) take note of the ethnomusicologist Nettle's conception of music as human sound communication outside the scope of language. Nettl describes the following musical elements that all cultures have in common: "song and dance; the making of some form of internal repetition and variation in their musical utterance; the use of rhythmic structures based on distinctions between note lengths and dynamic stresses" (in Mithen 2005:12). Morley (2003:3) adds Nettl's observation that the music of "all societies uses only three or four pitches, usually combining major seconds and minor thirds". He also refers to music as a non-referential communication system that, because of its profound emotional impact, is manipulative. These descriptions of music are effectively encapsulated Cross' (2003:2) conception of music: "Music embodies, entrains and transposably intentionalises time in sound and action" (see also Morley 2003:4).

From the discussions above, it is clear that music involves more than organised sound - it incorporates dance. Cross (2003:2) emphasises that "…. music is here not (easily) differentiable from dance; it might be that music and dance are simply two sides of the same coin, a view supportable by reference to some cultural practices outside those of the west which do not differentiate between 'music' and 'dance'…". The notion that music encompasses action has also been pointed out by Blacking (1973:27) who explained that for example "Venda music is founded not on melody, but on a rhythmical stirring of the whole body of which singing is but one extension". Mithen also remarks that a frequently ignored aspect of music is bodily entrainment. It

`

is artificial to separate rhythmic and melodic sound from rhythmic and melodic movement – song from dance. Therefore, in Mithen's work, "music encompasses both sound and movement" (Mithen 2005:15). Like Mithen, Morley emphasises that his definition of music incorporates both bodily and auditory elements of musical performance and perception for emotional communication. Musical movement is different from the many types of rhythmical movements that occur in everyday life. For example, a child rhythmically rocking on all fours or trembling caused by excessive excitation of the nervous system do not constitute musical movement. It is only motor behaviour that is intentionally disciplined that can be regarded as musical movement, or 'dance' (Hanna 1987). Hanna (1987:19-21) explains that dance is a human behavior composed of purposeful, intentionally rhythmical, culturally patterned sequences of nonverbal body movements other than ordinary motor activities. The purpose of dance is open-ended. In this study the production of organised sound and dance are considered inter-dependent or integrated musical actions.

These conceptions of music are compatible with Small's (1998) idea that music constitutes a deliberate action. This is why Small uses the term 'musicking' to describe music and dance and everything connected to it. He (1998:2) goes as far as saying that there is no such *thing* as music – music is not an object or a thing, but something that people do. The meaning, or significance of music does not only reside in the acoustic traces produced, but in the totality of musical action. This understanding of music is essential when studying the evolutionary origins of music because it directs the investigation towards the biological adaptations involved in executing musical actions.

As this study focus on the biological correlates for music, the development of musical instruments will not be discussed. The earliest evidence for musical instruments dates to 35 000 years ago (D'Errico *et al* 2003; Morley 2003) while the adaptations for song and rhythmical movement date to much earlier, around 1.6million years ago, as will be discussed in the following Chapters. In the following Section 2.3 the two elements of music, pitch and rhythm, are described and discussed. Even though timbre is also a musical element, there was insufficient information to incorporate this element into this investigation of the origins of music.

`

2.3 THE ELEMENTS OF MUSIC

2.3.1 Pitch

Pitch is the most frequently discussed musical element in music studies. The underlying physical dimension of pitch is frequency (Krumhansl 2000) as it is the frequency or rate of vibration of a physical source, for example a column of air or a string, which determines the pitch of a sound. The vibration results in a sound pressure wave that repeats over time. The number of cycles that the wave forms per second is known as the fundamental frequency (F_0) and this is measured in Hertz. The slowest vibration rate, or lowest sound would be the fundamental frequency and the associated sounds are the overtones. A musical sound, produced by a vibratory source like an instrument or the vocal chords consists of many harmonically related frequency components or partials. The fundamental of the note 'A' is 440 cycles per second. The second, third and fourth harmonic of 'A' are two, four and eight times this frequency. Even if the fundamental is removed the sound is still perceived as 'A' (Levitin 2006). Pure tones, or sine waves, are produced in a laboratory and these consist only of a single harmonic. In pure tones the air pressure rises and falls sinusoidally with time (Pierce 1996). Such tones are used in experiments with human and non-human subjects to test their perception and processing of sound.

Music relies on a discrete set of pitches, or a scale. From this scale melodies are constructed by choosing certain notes (Fitch 2006a:179). In musical melody the contour is defined by pitch direction and interval by the frequency ratios between successive notes (Peretz & Zatorre 2005:92). The pentatonic and diatonic scales are the most frequently used scales in all music (Carterette and Kendall 1999). There is a preference for intervals of fourths and fifths and for passages with a clear tonal orientation towards a 'root' in music (Thomson 2004:439). Levitin (2006) describes intervals of for example a unison, octave, and perfect fourth and fifth as consonant intervals.

2.3.2 Rhythm

In rhythm it is the time between events that is of most interest. Krumhansl (2000:161) notes that patterns of duration, rather than absolute durations have primacy, even though small differences in temporal duration can be discriminated accurately. Two types of time relations are basic to rhythm, the segmentation of an ongoing sequence into temporal groups based on their duration values and the extraction of an underlying regularity, known as beat or pulse (Krumhansl 2000). The temporal elements can be ordered in a regular or irregular way. A sequence, in which elements are ordered

`

regularly, as in the sound of a ticking metronome, is known as isochronous (Fraisse 1982).   Meter is "…the regular alternation of accents with one or more weak beats in a periodic pattern (corresponding to the bar)" (Krumhansl 2000:162). A musical beat thus occurs in the context of meter (Patel 2006: 100). Drake & Bertrand (2001:20-25) idenfity the following 5 universals of rhythmic processing in humans:

-    *Temporal grouping*

All humans group events that occur close in time together according to Gestalt principles in terms of characteristics such as timbre, pitch, intensity, duration and pauses. Even if there is no break in a sequence of events a change in any of these characteristics will be perceived as a break in the sequence and this leads to the creation of groups.

-    *Preference for regular sequences*

Each new event is compared with previous events and new stimuli are perceived as similar to previous stimuli within a "tolerance window". The perceptive system codes events in a relative manner as same/different, or same/longer/shorter.   Therefore irregular sequences tend to be heard as regular.

-    *Optimal processing at a rate of 600 ms*

Humans search for temporal regularities at a particular rate. There is an optimal processing zone and sensitivity to change is highest if events occur every 600 ms (1 ms is $1/10^{th}$ of a second) (Krumhansl 2001:160). The range varies from 300 – 800 ms interonset interval (IOI).   Many physical activities, from the beating of the heart to rocking and walking occur in periods of about 500 ms to 1 second and spontaneous tempo (natural tapping rate) varies from 300 to 880 ms (Krumhansl 2000). The average preferred or natural tempo for activities tends to center around 600 ms.

-    *Preference for binary rhythms*

 Humans have a natural tendency to perceive and produce longer intervals twice as long or short as previous intervals. This categorization principle favours binary, rather than ternary or more complex ratios (Bertrand & Drake 2001:25). This is termed subjective rhythmization (Krumhansl 2000:161). Infants detect temporal changes more readily in the context of strongly metric rhythms (Bergeson & Trehub 2006). This may be related to a presumed induction of an internal clock that facilitates

`

encoding of the temporal sequence, which leads to the detection of subtle (100 ms) changes. A fundamental pulse (Krumhansl 2000:173) of two events per second exists in musical and non-musical behaviours and this suggests that "a common internal oscillator may govern a variety of behaviours" (Krumhansl 2000:173).

- *Temporal regularities and synchronisation*

People spontaneously look for temporal regularities and organize activities around perceived regularities in all kinds of events. The rate and rhythm of everyday events are used to direct attention on a moment-to-moment basis through attentional synchrony (references in McAuly 2006:350). That is why humans are predisposed to find a regular pulse and synchronise with musical sequences. Humans find it difficult to break synchrony (Krumhansl 2000:160).

Sensorimotor synchronisation (Repp 2005) and entrainment or the movement of the body in synchrony to music are unique to human music making (Merker 2000; Repp 2005; Mithen 2005; Bispham 2006b). Entrainment involves reference to an external pulse as well as creating and controlling an internal pulse at will (Drake & Bertrand 2001). The ability to predict where the next beat will fall allows humans to synchronize their behaviour with that of the pulse (Merker 2000:316). Sensorimotor synchronisation thus takes place when an isochronous pulse is used to create "periodic temporal expectancies" that serve as the basis for motor synchronization to the beat (Patel 2006: 100). Humans engage in sensimotor synchronisation over a wide range of tempi. The range of accurate synchronisation is from 140 ms to 1600 ms (Krumhansl 2000:161). Sensorimotor synchronisation is used in many everyday tasks, but is fundamental in music performance of groups and dance (Repp 2005).

As will be discussed below and in Chapter 3, the majority of the adaptations relating to music that leaves traces in the fossil record can be associated with rhythmical movement. This makes rhythmical expression through dance of particular relevance here. Dance, as noted above, is purposeful, intentionally rhythmical, culturally patterned sequences of nonverbal body movements. The time expressed through motoric movement in dance is characterised by three elements, accent, duration and tempo. Accent is the relative force or intensity with which energy is relased; duration is the relative amount of time taken up by movements or groups of movements; tempo is the rate at which movements follow one another (Hanna 1987:30).

`

There is some debate on how rhythmic movements are generated. The timekeeper approach suggests that timekeepers generate a motor command after which a motor response is generated. The timekeeper generates the required sequence of responses. According to the nonlinear oscillator model (Beek *et al* 2000) preferred here, rhythmic movement and the associated physiological processes are manifestations of dynamic pattern formation or self-organisation. This self-organisation is in keeping with the dynamic systems approach. This principle has been used by Benzon, as will be discussed in Chapter 6, to argue for a link between rhythmic movement and entrainment of brain rhythms.

In Section 2.4 the biological apparatus associated with pitch and rhythm are discussed. The adaptations that make it possible to produce song in the form of controlled pitched vocalisation involve the vocal tract, respiratory and hearing apparatus (Sections 2.4.1 – 2.4.3). The ability to perform purposeful, intentionally rhythmical body movements that are an integral part of music is made possible by adaptations related to the inner ear and habitual bipedalism (Section 2.4.4.).

2.4 THE BIOLOGICAL ADAPTATIONS ASSOCIATED WITH MUSIC

2.4.1 The larynx and vocal tract

Singing is produced by means of the exhalation of air from the lungs that drives oscillations of the vocal folds or chord in the larynx or voice box. The rate of vibration of the vocal chords determines the pitch of a sound (Pierce 1996). The sound passes through the pharyngeal, oral and nasal cavities, collectively known as the vocal tract. In humans, the larynx is lower than in other primates and therefore the vocal tract is elongated. This elongated vocal tract permits a wider range of sounds. In other primates the high position of the larynx allows simultaneous breathing and swallowing. The larynx is initially in this high position in human infants, but between three months and four years of age the larynx descends to its adult position. In addition, it is out of the ordinary that humans are the only primates without a laryngeal air sac. All primates have inflatable air pouches that extend from the larynx and beneath the skin of the neck and thorax (Fitch 2000:260). It is not known how this affects vocal production, but it is a phenomenon that needs explanation.

Singing is a "highly evolved, uniquely human ability" (Perry *et al* 1999:3979). It involves the conscious and voluntary control of fundamental frequency, mainly through the vocal folds (Perry *et al* 1999). Musical singing varies from very simple chants and laments (see Avorgbedor 2008) to multifaceted melodies. Singing uses an open vocal tract (Frayer & Nicolay 2000:232). There are

`

differences in the way in which speech and singing is produced. In singing the entire laryngeal column moves up and down in relation to the pitch – a high larynx relates to high pitches and a low larynx to low pitches. In contrast, the position of the larynx in high speech sounds is inverted in comparison with singing (Vilkman *et al* 1996:83). Evolutinary investigation into the origins of singing needs to take into account that laryngeal biomechanics is complex and much research remains to be undertaken to understand the physiology of the voice (Vilkman *et al* 1996: 79).

The origins of human singing are conventionally discussed as a development from primate calls (e.g. Brown 2000; Geissman 2000; Morley 2003; Mithen 2005). However, Fitch (2006:183) contends that, if song is defined as complex learned vocalizations, there are no primates that sing. It needs to be demonstrated, rather than assumed that primate calls lead to singing. The possibility cannot be excluded that a simple laryngeal vocalization, termed a grunt, may have been the basis for the development of singing or complex learned vocalizations in hominins. Grunts in humans and a large number of nonhuman primates, including chimpanzees, gorillas, vervet monkeys and baboons are exerted under conditions of respiratory challenge, effort and locomotion (McCune *et al* 1996). A vocalized autonomic grunt occurs as a result of complex physiological processes. Under metabolic demand, when more oxygen is required, the intercostal muscles are activated to maintain lung inflation during expiration. This sets in motion a reflex contraction of laryngeal muscles that creates a system under pressure that lengthens the expiration phase of the breath and enhances oxygenation of the blood. Because the expiration is done against a constricted glottis, pulses of sound are produced. This is audible as grunts in humans and some larger animals, and is ultrasonically audible in small rodents. The grunt vocalization that follows from such autonomic constriction of the larynx is known as laryngeal breaking. Animals like chimpanzees, gorillas and vervet monkeys produce grunted vocalizations before they travel or when they observe others that are travelling (references in McCune *et al* 1996:28-30). Effort grunts are produced for example when infant chimpanzees climb over the mother's body. Vervet monkeys have different types of grunts that have referential function. One of the unique properties of human song is that it is learnt and complex. As grunts have been subjected to learning in other primates it would have been available to be acted upon in the evolution of singing. How grunts relate to learning is further discussed in Chapter 5, Section 5.5.Grunts, in addition to complex vocalization in apes and its relationship to the biomechanics of the larynx ought to be investigated in more detail in the context of the origins of music. In Chapter 3 the fossil criteria that have been used to infer the position of the larynx,

`

including bipedalism, the flexion of the basicranium and hyoid morphology (Morley 2002; Mithen 2005) are discussed.

A crucial difference between humans and other primates is the ability to voluntarily control the structure and complexity of vocalizations via the laryngeal and orofacial muscles (muscles that relate to the mouth and face) (Deacon 2000). The neural pathway responsible for controlling these muscles involves the periaqueductal gray matter (PAG) of the midbrain and the nucleus ambiguus (Morley 2003:123). Anatomically modern humans have a direct connection from the primary motor cortex to the nucleus ambiguus. Monkeys (and probably no other non-human mammal) do not have this direct connection (Jürgens 1992 in Morley 2003:122). Modern humans thus have direct control over the site of the laryngeal motorneurons and this allows them to have a larger degree of voluntary control over the structure and complexity of vocal utterances. Whether it is known when the changing neural pathways that allow control of the larynx evolved, is discussed in Chapter 4.

2.4.2 Respiratory control

Human singing would not be possible without the advanced degree of respiratory control that has evolved in the human lineage. Humans are unique in that they are able to control their breathing voluntarily (MacLarnon 1993; MacLarnon & Hewitt 2004). They have the ability to produce several sounds on exhalations enabling them to for example laugh by modulating upon a single out-breath. Chimpanzees use a sequence of repeated inspirations and expirations (Deacon 1997; Provine 2004). The capability for extended exhalation result from increased control of the release of air that involves the muscles that surround the lungs, the intercostals muscles and some abdominal muscles (MacLarnon & Hewitt 2004:182). This ability to control the breath is also used to modulate the volume and pitch of the sound to a much larger extent than other primates. Voluntary control of respiration is responsible for the fact that humans do not drop the fundamental frequency or pitch during sound production like non-human primates do (MacLarnon & Hewitt 2004:184).

A fossil indicator for the changed ability to control respiration and vocalisation is the size of the thoracic vertebral canal, as discussed in Chapter 3. As Fitch (2006a:196) explains, the motor neurons that control some respiratory muscles (which include the intercostals and abdominals) occur in this area, and therefore an enlargement in the thoracic area may indicate greater control over breathing. The research of MacLarnon & Hewitt (2004) referred to above was undertaken in

`

the context of speech, but the results are "equally, if not more, relevant to song" (Fitch 2006a:196). Singing requires a greater control of airflow than speaking (Sundberg 1987 in Fitch 2006a; Skoyles 2000; Frayer & Nicolay 2001:232). Fitch (2006a: 196) clarifies why singing requires finer respiratory control than speech: Singing uses lung capacity to a much larger extent than speech and singing requires the use of all the major respiratory muscles while speech requires only one set of intercostals.

2.4.3. Hearing and the outer, middle and inner ears

The production of song is crucially interrelated to hearing. Titze (1995) explains that the accurate perception of the pitch of your own voice is required to maintain an intended pitch through a process of corrective auditory feedback. Errors between the intended pitch and perceived pitch are corrected by muscle adjustments. Hearing involves the outer ear, middle ear, cochlea and labyrinthine systems of the inner ear. The outer ears or the pinna are responsible for the ability to judge where sound comes from (the height of the sound) (Pierce 1996). Sound is directed through the auditory canal or meatus, the eardrum or tympanic membrane to the middle ear, which consists of three small bones, known as the ossicles. The three bones of the middle ear, the hammer (malleus), anvil (incus) and the stirrup (stapes) are flexibly connected together in an air-filled space. The middle ear's main function is to amplify sound before it reaches the inner ear. The middle ear conveys the sound vibration through the oval window into the inner ear that is filled with fluid (Pierce 1996).

The inner ear consists of two interconnected parts – the cochlea and the vestibular system inside the bony labyrinth. The cochlea is the main organ of hearing as it is involved in the analysis of the vibratory spectrum of the sound. The organ of Corti is situated on the basilar membrane of the cochlea. The hair cells of the organ of Corti translate the sound vibrations into electrical impulses that are relayed to the brain via the auditory nerve. The auditory nerve conveys the electrical impulses to the brain (Moggi-Cecchi & Collard 2002). Only certain hair cells fire in response to certain frequencies. Low frequency sounds excite hair cells on the one end of the membrane, while high frequency sound excites the cells at the other end. The association of certain pitches that excite certain specific areas of the membrane is referred to as a tonotopic map. If the membrane is activated, electrical signals are sent to another tonotopic map in the auditory cortex. Pitch, unlike almost any other musical attribute, is represented directly in the brain (Levitin 2006:27). The bones

`

of the inner ear do fossilize and can potentially inform on the evolution of hearning. The functioning of the vestibular system is further discussed below in 2.4.4.

2.4.4   Rhythmical bodily movement

In comparison to the literature on the production and perception of song, there is little information on the production of rhythmical body movement in dance. This may be because the study of moving bodies has been of peripheral interest in anthropology (Reed 1998). Dancing involves the volitional rhythmical movement of the whole body interacting with gravity (Hanna 1987). It uses the large muscle groups to voluntarily move the extremities in a sensorimotoric-synchronised fashion. The ability to dance is related to habitual bipedalism. Habitually standing or walking on two legs requires the constant use of muscle groups to monitor the centre of gravity and to integrate the movement of the legs with the arms, hands and trunk (Clarke 2005:66). Imbalance in this system results in Parkinson's disease. Playing music to patients with Parkinson's disease sometimes relieves their motor difficulties and it is thought that it is the rhythmical content of the music that has this effect. In this regard Morley (2003:211) discusses the work of Thaut *et al* (1997) who showed that an external rhythmic stimulus plays an important role in gait control - the auditory rhythm entrain motor patterns.

The shift to bipedalism may have been crucial in the development of rhythmic entrainment and beat induction. The process of beat induction, or "the activation of a regular isochronous pattern (the beat) when listening to regular temporal sequences" is one of the key elements in rhythm perception (Todd *et l* 2007:1). It has been suggested that musical rhythm may originate in part from the motor rhythms controlling locomotion (Trainor 2008:598). Phillips-Silver & Trainor (2005, 2007, 2008) have undertaken several experiments to test the relationship between movement and rhythmic perception and production. They demonstrate that the way in which infants and adults move their bodies, (including the head) to music affects their auditory perception of rhythm structure. For example, in their 2005 study it is reported that whether adults and infants are bounced on the 2nd or 3rd beat of a repeating 6-beat rhythm pattern with no accents influence them hearing or preferring a march or waltz respectively (Train 2008). This, according to Trainor (2007:18) "presupposes the ability to perceive different tempos and to entrain to different tempos. It also presupposes the ability to perceptually group sound events into a rhythmic hierarchy.".

`

Clark (2005:66) points out that the ventromedial muscles that control the larger movements of the body interact with the vestibular apparatus of the inner ear (Clarke 2005:66). The vestibular system consists of the vestibule and semicircular canals (Spoor *et al* 2007). This structure is important for stabilizing the gaze during locomotion (Spoor *et al* 2007) and for regulating balance. It plays a role in coordinating upright bipedal behaviour through monitoring body movements in the vertical plane. It is evident that hearing and movement are closely interrelated. Sufficiently loud sound affects the balance sensors of the vestibular system (Carey & Amin 2006:482).

The vestibular system plays a pivotal role in the interaction between movement and the perception of musical rhythm (Todd & Lee 2007:111; Phillips-Silver & Trainor 2005:1430). Phillips-Silver & Trainor (2008:94) found that "…metrical encoding rhythm can be biased by passive motion. Furthermore, because movement of the head alone affected auditory encoding whereas movement of the legs alone did not, we propose that vestibular input may play a key role in the effect of movement on auditory rhythm processing". The important inference for the origins of music is that the vestibular system is involved in auditory rhythmic processing. The vestibular structure of the inner ear fossilizes and therefor provides an opportunity to study the evolution of movement and rhytmical processing as further discussed in Chapters 3 and 5.

2.5 DISCUSSION

The intention of this Chapter was to construct the analytical framework for the investigation into the origins of music. The definition of music and the musical elements of pitch and rhythm were discussed. "Music" has been defined here as an intentional action in which complex, learned vocalizations (and/or instrumentally produced sound) are combined with the movement of the body in synchrony to a beat. Whole body movements are involved.

The biological adaptations relevant to the production of musical pitch and rhythm have been described. The biological adaptations underlying musical expression and perception include a lowered larynx and voluntary control of the respiratory muscles. This, together with uniquely evolved brain mechanisms, allows humans to produce a larger variety of sounds and to control the fundamental frequency and dynamics of the sounds. The evolution of humans' middle and inner ear can be related to their hearing capabilities and their ability to control vocalizations. The crucial adaptations that can be linked to the ability to move rhythmically are related to bipedalism and the vestibular system of the inner ear. In Chapter 3 the fossil evidence for the descent of the larynx,

`

respiratory control, the middle and inner ear and bipedalism as discussed by Mithen and Morley is presented. Their propositions are critically discussed at the hand of relevant literature and further suggestions are made for the timing of the appearance of these crucial adaptations. Chapter 3 also engages with the context from which this study is undertaken by critically discussing the musilanguage hypothesis.

`

CHAPTER 3: A CRITIQUE OF THE CONCEPT OF MUSILANGUAGE AND FOSSIL EVIDENCE FOR THE MUSICAL CAPABILITIES

3.1 INTRODUCTION

Mithen's and Morley's investigations of musilanguage respectively incorporate information from disciplines such as anthropology, archaeology, palaeoanthropology, neurology and developmental psychology. Their syntheses discuss similar concepts, but Mithen sets out to explain how a musilanguage could have developed into language and music, whilst Morley investigates the evolution of music only. Morley and Mithen propose that the intermediate stage between primate affective vocalization and fully syntactical language was a 'musilanguage' or Hmmmmm. Musilanguage incorporated voluntarily expressed short melodic phrases that were not symbolically referential accompanied by rhythmical synchronized body movements and gestures. These melodic phrases were learnt via complex vocal learning and contained new or novel material. The early holistic phrases would have made "extensive use of variation in pitch, rhythm and melody to communicate, express and induce emotion" (Mithen 2006:98). Similarly, Morley (2003:181-183) argues that musilanguage's increased range and control of pitch allowed greater vocal versatility and expressiveness in vocal affective communication. Interdependent with the prosodically-contoured vocalisations would have been corporeal and rhythmic expression of affect. These elements formed the foundations of musical behaviours. The aim of this chapter is to present and appraise their theories in terms of primate origins, structure, evolutionary scenarios and fossil evidence.

3.2 THE PRIMATE ROOTS OF MUSILANGUAGE

It is widely thought that primate affective animal vocalization contained the "building blocks" for music (Richman 1993; 2000; Scherer 1991, 1995; Papoušek , H. 1996; Vaneechoutte & Skoyles, 1998; Brown 2000; Huron 2001; Merker 2001; Tolbert 2001; Morley 2002, 2003; Winkelman 2002; Cross 2003; Christensen-Dalsgaard 2004; Fitch 2005a, 2006a,b; Mithen 2005). The majority of these authors discuss music in the same evolutionary framework as language. Continuists such as

`

these experts regard primate calls as suitable evolutionary precursors for music as well as language. The issue is whether primate calls (or other primate behaviours) could have been homologues or analogues for musical behaviours. Homologues are traits that are similar in two or more species because they derived from a common ancestor. The trait may be in a somewhat different form or have a different function (Fitch 2005a). By contrast analogues are similar behaviours or similar solutions to problems that developed independently in two different evolutionary lineages.

Mithen (2005a:38) mentions that the vocalizations, gestures and body postures used by non-human primates are probably analogous to those used by early hominins. The musicality of primate vocalisations reside in for example the rhythmic chattering of geladas and the duets of gibbons. Morley (2003) discusses the relationship between complex vocalisations of primates and humans, but does not explicitly state that primate calls were homologous or analogous to a musical protolanguage. Animal vocalisation can only qualify as 'song' comparable to human musical singing if it is learned and complex (Fitch 2006a:178). Vocal learning depends on the ability to imitate novel sounds (Fitch 2005a:35) and it is not clear what 'complexity' refers to. Fitch notes that the various innate vocalizations in humans (groans, sobs, laughter and shouts) are not complex. "Complexity", however is a vague standard, because, as Fitch (2006a:178) remarks, there is no widely accepted metric for complexity applicable to all musics.

There is abundant comparative evidence (Fitch 2006a) that complex, learned vocalization or song evolves relatively easily and therefore 'deep' similarities between human and animal song is highly significant for the evolution of music (Fitch 2005a:35). Song-like vocalizations have been discovered in cetaceans (whales and dolphins) and the pinnipeds (seals and sea lions), gibbons and suboscine birds, but not, apparently in primates. According to Fitch (2005a) the vocalizations of gibbons and suboscine birds are not song-like because they do not depend on vocal learning (Fitch 2005a). He (Fitch 2005a:35) concludes that no nonhuman primates, including the apes have the ability to learn novel sounds because of fundamental differences in the neural control of vocalization: " Despite some similarities in form and function…., and clear homology at the level of the vocal production system, the lack of extensibility of primate calls renders them categorically different from human music and speech. This difference between humans and other primates appears to be underlain by fundamental differences in the neural control of vocalization". Fitch therefore inferred that primate vocalizations, even though pitched and complex, are not relevant to the origins of music (see also Marler 2000).

`

This conclusion may be premature because Fitch's conclusion on primates' inability to learn novel vocalizations is based on less than a hand full of studies. The neural control of the larynx in apes and humans are similar, although humans have more extensive control (Deacon 1992). There are too few long-term studies on ape calls in the wild to claim that their vocalizations are not learned. Moreover, there is a degree of flexibility in ape vocalizations in terms of "temporal organization of the frequency modulation of their tonal elements, particularly F_o" (Masataka 2007:36). Masataka (2007:37) describes long-distance calls of apes as typified by "pure tonal notes, stereotyped phrases, biphasic notes, accelerando in note rhythm, and possibly a slow-down near the end of the phrase. In Chapter 2 it has been surmised that vocal learning could have evolved through grunts that could have been combined with rhythmical movements. It is conceivable that ape-call-like vocalisations were part of hominins' vocal repertoire and that learning obtained through grunt vocalizations could be transferred to such calls to produce the first song.

Continuists conventionally focus on whether human song evolved from primate song, but other aspects of musical behaviour may be homologous in apes and humans. Fitch (2005a) suggests that percussive behaviour or drumming may be homologous in apes and humans because bimanual percussion on resonant objects commonly occur in chimpanzees, bonobos and gorillas, but rarely in other species (Fitch 2005a:37). Chimpanzees further display behaviour that may be analogous or homologous to dance. They commonly take a bipedal stance, stamp their feet rhythmically, and make hooting noises, together with arm and facial gestures (McNeill 1995). Jane Goodall (in McNeill 1995:16) observed that chimpanzees sometimes react to a thunderstorm with a group dance-like display (the movements were entrained). McNeill (1995:17) also discusses Kohler's description of how chimpanzees would spontaneously engage in entrained movement whilst moving around a post in their compound. This is significant, notwithstanding the fact that Jane Goodall and Wolfgang Kohler did not observe the chimpanzees in a completely wild environment.

From this short discussion it is evident that there is a need for increased empirical study into ape 'musical' expressions with the purpose of researching musical correlates. Aspects like learned vocalization and entrained percussive and rhythmical movement must be investigated in more detail before the hypothesis that ape vocalization and other rhythmical and entrained movements are homologues for human musical expression can be rejected or accepted.

`

3.3 THE STRUCTURAL PROPERTIES OF MUSILANGUAGE

The structure of musilanguage is discussed here to consider the advisability of investigating the origins of music and language within a single framework. Mithen and Morley uncritically accept the structure of Brown's (2000) "convincingly argued" (Morley 203:144) three-stage model for musilanguage. Brown (2000:271) sees the structural properties of musilanguage as an outgrowth of homologous precursor functions. In both language and music a phrase is the basic unit of structure and function. Phrases are generated from a limited repertoire of discrete units and use basic acoustic properties to convey emphasis and emotional state.

According to Brown fully-fledged syntactical language developed from the three common features of music and language - lexical tone, combinatorial syntax and expressive intonation. The first stage in the evolution of musilanguage would have involved lexical tone. "Lexical tone, with its underlying level tones and semantically meaningful pitch movement would have been a joint feature of language and music and a scaffold on which both systems could have developed" (Brown 2000: 285). Lexical tone would have used discrete pitch levels (level tones) to convey (semantic) meaning. Brown supports this by the evidence from autosegmental models in which phonological processes, such as tone and vowel harmony are independent of and extend beyond individual consonants and vowels. However, changes in pitch levels in musical expressions (including prosody) do not communicate meaning or semantics in a straightforward universal way (Cook 2001; Shepard & Wicke 1997) - the consensus is that musical expression is not referential. There seems to be insufficient evidence to argue that pitch levels could have conveyed meaning.

In the next 'combinatorial' phase, lexical-tonal units would have been strung together to form tonal contours. These are unordered phrases with 'higher-order' meaning (Brown 2000: 285) that convey emotive and/or pragmatic meaning. The contour of the phrase determines the emotional meaning of the phrase. Surprise and question intonations are examples of such phrases. These phrases would have had a rhythmic structure as well, derived, in part, from the temporal arrangement of the elemental units. The third level involves expressive phrasing. Intonational phrasing occurs via expressive properties such as tempo, dynamics and rhythmic modulation for the expression of emotion and emphasis. Intensity of emotion is expressed via tempo modulation (slow-fast), amplitude modulation (soft-loud) and register selection (low-high). For Brown this system of human emotional expression is universal because in speech, gesture and music the same sentic profile (Clynes 1978) is used to express a given emotion intensity state, regardless of the modality

`

of expression. "For example, happy music and happy speech are both characterized by fast tempos, large-amplitude sounds, and high registers; sad music and sad speech are characterized by the opposite sentic spectrum" (Brown 2000:288). The importance attached to emotion in the discussions of musilanguage by Mithen and Morley originates from this understanding of music.

The idea that "music" communicates "emotion" is contested (e.g. Scherer 2003; Shepard & Wicke 1997). It has never been deomonstrated that a certain component of music has a certain specific psychophysisological effect that can be tied to a certain 'basic' emotion (Panksepp & Panksepp 2000). Neither has it been shown that these elements are the same in prosody and music. There is very little evidence to support Brown's (2000, 2006) argument that the precursor to language and music involved a sophisticated referential *emotive* system in which broad semantic meaning preceded precise semantic meaning. The evolutionary relationship between "music" and "emotion" is further discussed in Chapter 6.

Mithen not only draws on Brown's work, but also on Wray's (1998; 2000) proposed structure of protolanguage. Alison Wray's structure of protolanguage has similarities to Brown's musilanguage although she does not explicitly tie it to musical expression (Wray 2006). Her view of protolanguage (Wray 1998, 2000; Wray & Grace 2007) is the reason for the holistic and manipulative in Mithen's Hmmmmm. She suggests that holistic phrases, which could be said to be roughly comparable to Brown's unordered phrases with 'higher-order' meaning, preceded linguistic syntax. Holistic phrases could have consisted of a string of syllables such as 'tebima'. This may have had a holistic meaning just like 'abracadabra' today may be translated as 'I hearby invoke magic' (Mithen 2005:149). Holistic utterances would not have mapped onto specific entities or actions and were not composed of discrete entities or 'words' (Mithen 2006). An example of holistic and manipulative calls are those of vervet monkeys. These calls are complete messages and are holistic because they have no internal structure and have not been combined with any other vocalization. These types of alarm calls are manipulative rather than referential because monkeys are not telling their group members about the world; they aim to manipulate their behaviour (Mithen 2005:120). Further examples of holistic phrases are generic forms of greetings, statements and requests (Wray 1998; Mithen 2005:172). Bickerton (2007) and Tallerman (2007) are critics of the holistic view of protolanguage. Bickerton (2007) cannot see how 'holophrasic' or holistic phrases could be segmented and sees protolanguage as "containing a categorically complete, if severely limited vocabulary of items roughly equivalent to modern words, but lacking a

`

sophisticated phonology and any consistent structure" (Bickerton 2007:517). Thus words came before grammar. For Tallerman (2006, 2007) ape-like vocalization is very different from human language and cannot be a precursor to modern language. Tallerman (2006, 2007) argues that different parts of the brain relate to speech and primate vocalizations, that primate calls and human speech use different physiological bases, that primate calls are genetically transmitted, that the dissociation between sound and meaning in human speech is absent from primate calls and that primate calls are involuntary whereas speech is voluntary. In Mithen's (2005: 254) response to Tallerman's criticisms he mentions that some of these objections are based on 'misconceptions' of primate vocalization and holistic protolanguage. It is beyond the scope of this study to evaluate these objections to the holistic view of protolanguage. However, the objections show that the way in which Mithen and Morley link musilanguage and grammatical language is problematical. The argument that musilanguage could have lead to modern language needs careful and in-depth study in its own right. Botha (*in press*) undertook such an exercise and found no support for the argument that Hmmmmm could have been a precursor to grammatical language. This is an important reason for investigating the evolutionary origins of music separately from the evolutionary origins of grammatical language.

3.4 EVOLUTIONARY RATIONALES

In biology, answering the 'why ' type of question is the hardest – the evolutionary function of a trait often changes over time and could be exapted for another purpose (Fitch 2006). Multiple selective routes to musical expression may have existed. In considering the reasons for the evolution of a musical protolanguage, Mithen and Morley cover similar ground. They discuss sexual selection, group cohesion and infant directed communication as possible selective agents.

3.4.1 Sexual selection

It is often suggested that the function of music was to attract sexual mates (e.g. Darwin 1871; Miller 2000, 2001). In sexual selection scenarios the inheritance of desirable traits through the choice of a mate ensures reproductive success. Miller (2000, 2001) is a well-known proponent of the idea that singing and dancing were used in the Palaeolithic to advertise sexual fitness. Mithen (2005:179-180) and Morley (2003:190-193) convincingly question the logic of his arguments. They also question Merker's version of musical expression and sexual selection. In Merker's (1999, 2001) view co-ordinated synchronous rhythmic and melodic group behaviour or synchronous chorusing

`

activities were carried out to attract female hominids from neighbouring territories. This proposal assumes that the Palaeolithic populations were exogamous (males stay in their natal group while females emigrate). Groups of males would have undertaken synchronous chorusing activities because coordinating voices carry further than one alone. However, Morley (2002:207) points out that there is no precedent for this type of co-operative behaviour amongst primates, as it is principally known amongst insects. Also, it implies two conflicting manifestations of the same behaviour – co-operative and then competitive. Such chorusing would also have attracted predators (Mithen 2005:207). Morley (2003) mentions that sexual selection is an unlikely rationale for the evolution of music because, in contrast to animal behaviour where sexual selection for song is mostly male behaviour, in humans both genders have musical abilities (Morley 2003; see also Fitch 2006). More research needs to be undertaken to address the shortcomings of the sexual selection hypothesis before it can be applied to the origins of music. For this reason Mithen's (2005:189-191) idea that Hmmmmm could have been used to advertise sexual availability, especially when combined with making handaxes, can only be considered as a thought-provoking speculation.

3.4.2 Group cohesion

The hypothesis that musical behaviour may have functioned to strengthen social bonds with conspecifics to promote group-cohesion (Brown 2000; Morley 2003; Mithen 2005) is more credible than the sexual selection hypothesis. Mithen (2005:214-217) argues that those individuals who used Hmmmm to make music together would have suppressed their self-identity to work towards a group identity. The australopithecines, the first *Homo* and *Homo ergaster* would have had to live in large complex social groups so survive (Mithen 2005:217). To thrive, cooperation with others and trust between group members would have been necessary. He supports the ideas of McNeill (1995), Freeman (2000) and Benzon (2001) who have proposed that joint music making forges social bonds and group identity. Morley (2003:147) suggests in similar vein that an individual able to maintain a large cohesive social network by means of vocal agility to provide support in power, mate or food resource contests would have considerable selective advantages over its peers.

Mithen (2005:135-6) and Morley (2002: 212) relate the evolution of complex vocalization to the social grooming hypothesis (Aiello & Dunbar 1993; Dunbar 1998). According to this hypothesis complex vocalization became more important over the course of hominid evolution because group size increased to its current average of about 150 individuals. Aiello & Dunbar (1993) noted a proportional relationship between the degree of encephalization and group size amongst higher

`

primates and humans. Regressions were used to show the relationship between cranial capacity, group size and grooming time. With the enlargement of group size there would have been selective pressure for the utterances to become more efficient and expressive. Early vocalizations would first have been used to facilitate and then to replace social grooming. Morley (2002:210) remarks that these utterances may have been emotive vocalizations and Mithen speculates that the first emotive utterances may have been "closer to song" (Mithen 2005: 136). Even though Morley and Mithen express a degree of scepticism about the relationship between group size and brain size, they still regard it as a plausible explanation. However, the methodology used to infer the relationship between brain size and group size is flawed (Noble and Davidson 1997; Davidson 1999). There are no firm grounds for the argument that group size imposed the need for increased social grooming, whether spoken or sung (see also Byrne & Bates 2007).

The scenario that Mithen (2005:218-220) creates around the assemblage of *Homo heidelbergensis* remains at Sima de los Huesos (the cave of bones), Atapuerca in northern Spain, is therefore also unlikely to be true. The remains of at least thirty-two individuals, dating to 350 000 years ago, were found in this cave. It seems that complete bodies were deliberately dropped into the pit. Because of the similarity in the state of bone preservation, their density and frequently present bone articulations, the researchers (Bocquet-Appel & Arsuaga 1999) contend that these archaic humans died together or within a very short time of each other. Mithen (2005:220) speculates poetically that they would have used Hmmmmm to communicate about these events: "It would have been a time to consolidate and confirm social bonds, and to make a commitment to future cooperation. Communal Hmmmmm singing and dancing would have achieved this: sound and movement expressing grief and facilitating 'boundary loss' among the individuals who wept 350 0000 years ago in the cave at Atapuerca". There is no concrete evidence for this scenario – it cannot be used to support the notion of musically advanced *Homo heidelbergensis*.

The problems involved in determining prehistoric group size and correlating it with brain size mean that the social grooming hypothesis is not plausible. However, the hypothesis that musicking led to group cohesion and trust is plausible, even though it is unlikely that it could be determined which pressures led to an increased need for group cohesion and trust. It is important not to confuse this group function with group selection (Morley 2003; Fitch 2006) - although musical behaviours may have involved groups, individual selection and not group selection drove the evolutionary process. In Chapter 6 it is discussed how musical activities could have engendered trust and group-cohesion.

`

3.4.3 Infant-directed communication

Infants seem to have an innate preference for the melodic, slow and repetitive way in which caregivers communicate with them. They have the capacity to perceive melody and rhythm and to gain affective information from this communication. They rhythmically synchronise their movements and vocalisations with those of their caregivers (Malloch 1999; Trevarthen 1999). The question is whether the needs of infants could have been the incentive for the evolution of musilanguage. Dissanayake (2000) and Falk (2004) propose that Infant-directed communication evolved to deal with the extreme altriciality or helplessness of hominin infants. For them musicality is based on the capabilities that evolved within this context.

For Falk (2004) the root of Infant-directed communication or motherese is to be found in the behaviour of chimpanzees. Baby chimpanzees are unable to cling to their mothers in the first two months of their lives, and vocal communication is crucial in eliciting maternal support. Chimpanzee infants use vocalisations, including the whimpering call sequence, screams and hoo's. It is from this largely affective type of communication that "motherese evolved incrementally" (Falk 2004:497). The need for increased maternal care was created by the increasing altriciality of babies in the Homo lineage.   The evolution of bipedalism resulted in a smaller pelvis and as a consequence mothers gave birth to babies that are much more undeveloped than chimpanzee babies and are helpless or altricial for much longer. This means that the mother has to hold the baby all the time during activities. The hominin mother may have needed to put the baby down more often than chimpanzee mothers to forage effectively. She would have 'parked' the baby and reassured him using prosodic vocalisations to keep him from crying and fussing. This changed the mother-infant contact from being exclusively based on direct physical contact to being distal. 'Parking the infant' led to the development of infant-directed speech and lullabies. Melody preceded, and developed into semantic language using words (Falk 2004:501). But, as Burling (2004) points out, it is not explained satisfactorily how words could have emerged from prelinguistic melody (see also the commentary of King & Shanker, Bickerton, Bortfield, Bouissac and Burling in Falk 2004).

The argument of Falk and others who propose that the order of language acquisition in infants represents its evolution rests on the assumption that 'phylogeny recapitulates ontogeny', as Bouissac (2004) mentions. Haeckel, in 1866, proposed that ontogeny recapitulates phylogeny. The assumption was that the growth of an embryo (ontogeny) recapitulates or demonstrates the evolutionary history of a species (phylogeny). This idea has been severely criticized. Falk

`

(2004:527) however strongly disagrees that her theory assumes that 'ontogeny recapitules phylogeny' (ORP). Rather, it rests on the assumptions of the modified version of ORP known as 'altering ontogeny formulates new phylogeny' proposed by Goodman & Coughlin (2000). Determining in which respects this is different from ORP is beyond the scope of this study, but in principle, it cannot be assumed that the developmental history of an organism reflects its phylogenetic history. That melodic communication precedes the acquisition of words in infants does not indicate that melodic communication preceded and led to linguistic communication in the evolution of Homo.

Mithen (2005) proposes that Infant-directed communication could have facilitated infants' emotional development and could have provided selective pressure for the evolution of a multimodal emotional vocal protolanguage. Morley (2003:136) on the other hand points out that increased affective vocalization may equally apply to interaction between any individuals within a hominin group. He suggests that motherese can be regarded as a specialized perpetuation of a non-linguistic interaction that was earlier used between all individuals (Morley 203:222) rather than a model for the evolution of language. Morley's proposal is more likely Mithen's for the following reasons: Infant-directed communication is not unique to humans. It is common in mammals and is also found in some bird species (Newman 2004). Furthermore, maternal behaviour is one of the basic emotional operating systems (Panksepp 1998) that occur in many species. Therefore the existing hominin infant-directed communication system may have been sufficient to cope with the demands of altricial infants. The specific human expression of infant-directed communication probably reflects musicality evolved in another context. The significance of Infant-directed communication for the origins of music is further discussed in Chapter 5.

3.5 MODERN MUSICAL AND LINGUISTIC ABILITIES

The evolutionary process that led to modern musical and linguistic abilities was, according to Mithen, complete by 40 000 years ago. The subject of Mithen's (2005) book is the Neanderthals who lived between 200 000 and 35 000 years ago in Europe, and therefore they feature prominently in his discussion of Hmmmmm. Mithen argues that Neanderthals were musically advanced, but did not have modern language and speech abilities. They were "intensely emotional beings" and would have had an advanced form of Hmmmmm. Mithen describes Neanderthals as musical savants who evolved unique neural networks for the musical features of Hmmmmm not present in *Homo sapiens*.

`

Three features provide 'compelling evidence' (Mithen 2005:228) that this was the case: Neanderthals lived in small socially intimate communities (there would have been no need for compositional language), symbolic artifacts are absent and the continued presence of 'immense cultural stability'. This leads Mithen to propose that Neanderthals would have had a domain-specific intelligence that imposed major constraints on their levels of creativity and symbolic thought (Mithen 2006: 100). In this domain-specific mentality each type of intelligence had been moulded by natural selection to thought patterns and knowledge suited for solving problems within its relevant behavioral domain. The three key intelligences are social, natural history and technical intelligence. Neanderthals, for example would not have been able to transform a bone into a flute because this would have required taking an object from the natural history realm and modify it for use in the social realm and this requires cognitive fluidity (Mithen 2005: 243). They lacked the additional neural circuits that made connections between these domains, which may have required very little extra capacity (Mithen 2005:233).   Cross (1999; 2001) similarly argues that proto-musical behaviours served to integrate information across the physical and social domains and this would have given rise to a flexible, cross-domain intelligence or cognitive fluidity in modern humans. *Homo sapiens* mind would have had cognitive fluidity in which 'ways of thinking' and 'stores of knowledge' (Mithen 2005:263) were integrated to create new types of thoughts. This would have enabled mixing human and animal features in sculptures, for example and making musical instruments in animal bone. This modular model of the mind rests on the principles of Evolutionary Psychology (Panksepp & Panksepp 2000) that, as discussed in Chapter 4, are questionable on various grounds.

The 'compelling' evidence put forward by Mithen can be challenged – there is no archaeological proxy for prehistoric group size; there is evidence for symbolic artifacts associated with Neanderthals (d'Errico *et al* 2003) and the 'immense cultural stability' collapses under close scrutiny (e.g. Soressi 2005; Zilhão 2007). The evidence that Mithen discusses cannot be used to substantiate the hypothesis that Neanderthals were musically gifted but linguistically underdeveloped.

Morley (2003:222) is ambiguous about the point at which modern music, and syntactical language developed. He puts this development at some point between archaic *Homo sapiens* and anatomically modern *Homo sapiens.* He (2003: 220) does not see musical behaviour, and thus its emergence, as relying on the presence of a specific "music module" or the breaking down of

`

barriers between cognitive "domains of intelligence". One of the aims of this study is to investigate whether the biological and fossil evidence support Mithen or Morley's processes underlying and timing of the development of modern music as defined in Chapter 2. In Section 3.6 the fossil evidence for musical capabilities is discussed.

3.6 THE FOSSIL EVIDENCE FOR MUSICAL CAPABILITIES

3.6.1 The descent of the larynx and vocal production

In Chapter 2 it was discussed that the lowered larynx in humans makes a wider range of vocalizations possible. This may have been a key evolutionary event linked to musical vocalizations. It is generally argued that the selective cost of the descent of the larynx must have been great because it increases the risk of choking. Therefore its selective advantage must have outweighed the cost (Lieberman 2007). However Clegg & Aiello (2000 cited in Morley 2003: 85 and Mithen 2005:146) note that the cost of laryngeal lowering may have been exaggerated. Morley (2003) and Mithen (2005) both regard the descent of the larynx as pivotal to the evolutionary origins of music and both regard habitual bipedalism in *Homo ergaster* as the cause of its descent (Morley 2003:169; Mithen 2005:248). In this section the fossil indicators used by Mithen and Morley to infer the descent of the larynx are discussed. The relevance of these indicators is then evaluated at the hand of pertinent literature.

The descent of the larynx is often tied to the evolution of bipedalism. Bipedal locomotion has lead to changes in hominin anatomy in for example the length of the limbs, the way in which the hip, knee and ankle joints function, and the form of the toes (Mithen 2005:141). Mithen (2005:146) accepts the proposition of Aiello (1996a) that bipedalism caused the descent of the larynx. Aiello (1996a) proposes that walking upright required the spinal cord to enter the brain case from below rather than behind and this left less space for the larynx between the spinal cord and the mouth. This in conjunction with the reduction of the lower jaw in early hominins forced the larynx down the throat and incidentally the vocal tract became longer, allowing a greater range of vocal sounds.

Morley (2003:220) is more ambiguous about the relationship between bipedalism and the descent of the larynx. He suggests that the larynx' descent may have been a by-product of bipedalism or a consequence of advantages associated with increasing versatility and control over complex tonal affective vocalization (:149). Another suggestion is made by Fitch (2000) who argues that the selective force for the (initial) lowering in *Homo ergaster* of the larynx was to exaggerate size. The

`

increased vocal tract length that resulted from a lowered larynx decreased the formant frequencies and allowed individuals to exaggerate body size in dense foliage or at night. Morley (2003:128-131) reacts to this hypothesis by pointing out that at the time of the lowering of the larynx, *Homo ergaster* moved away from arboreal habitats and that the animals used by Fitch to demonstrate the link between the lowering of the larynx and exaggeration of body size temporarily, not permanently, lower the larynx. Morley's line of reasoning is convincing. However, Nishimura's work indicates that bipedalism may have had little to do with the descent of the larynx.

The research of Nishimura (2003, 2005, 2006; Nishimura *et al* 2003) provides a completely different perspective on the descent of the larynx. According to Nishumura's hypothesis the descent of the larynx was a two-step process and not limited to the human lineage. The "descent of the hyoid bone in relation to the mandible and cranial base, and that of the laryngeal framework in relation to the hyoid bone" affect the position of the larynx (Nishimura 2003:41). MRI images and studies of living chimpanzees were used to determine that the larynx descends (to the extent that it loses contact with the velum) in both chimpanzee and human infants. In humans, however, the descent is more pronounced. In humans there is an additional descent - the hyoid bone also descends relative to the palate. In Nishimura's (2003:47) words: "It is, therefore, proposed that evolution of the descent of the larynx was in fact a two-stage process: the descent of the thyroid followed by the descent of the hyoid."

That humans and chimpanzees share the first phase of descent means that the last common ancestor of chimpanzees and humans who lived before 6 million years ago must have shared this feature. This descent may be related to changes in the swallowing mechanism as an adaptation to changes in hominoid diet and increase in body size (Nishimura *et al* 2003: 6933). The second descent, that of the hyoid relative to the palate, has not been identified in non-human primates and must be a development restricted to the hominin lineage. This second descent in hominins was probably caused by modifications of the facial-mandibular skeleton in hominin evolution, for example the decreasing prognathism of the facial skeleton (prominence of the jaw) and increase in mandibular height (Nishimura *et al* 2003:6933). The descent of the larynx thus had little to do with bipedalism or the production of sound. The lower position of the larynx caused by the second descent, however, incidentally made the production of a wider range of sounds possible.

The question is whether there are any fossil indicators that may point to the timing of the second descent. Morley (2002: 198; 2003: 77-87) regards the flexion of the basicranium and hyoid

`

morphology as useful indicators of laryngeal position in fossils. However, these attributes do not indicate the position of the larynx: The research of Lieberman & McCarthy (1999), Fitch (2000) and Lieberman (2007) convincingly shows that the flexion of the basicranium is not related to the position of the larynx. Similarly, Fitch (2000:262) has demonstrated that hyoid morphology does not reliably indicate the position of the larynx. The most meaningful conclusion that can be reached from a much-discussed Neanderthal hyoid bone of anatomically modern configuration (Frayer & Nicolay 2001) is that it has no hollows into which air sacs could have fitted. Neanderthals have lost the air sacs that are common in primates (Fitch 2000:262).

The initial descent of the larynx predates the split between chimpanzees and the Homo lineage 6 million years ago and it is unknown when the second phase of descent took place. There are no clear-cut links between skeletal morphology and vocal tract anatomy and therefore the fossil record cannot conclusively inform on the descent of the larynx. This then makes it highly improbable that there was a bootstrapping or self perpetuating evolutionary effect between the lowering of the larynx postulated for *Homo ergaster* and the neurology that allow increasingly complex planned vocal sequences of the laryngeal and orofacial muscles (Morley 2003:123). There is insufficient evidence to argue, as Mithen and Morley do, that the descent of the larynx played a role in the evolutionary origins of music 1.6million years ago.

3.6.2 The size of the thoracic canal and respiratory control

Without elaborate respiratory control singing, as explained in Chapter 2, would not have been possible. The first development that influenced the way in which hominins control breathing was the evolution of bipedality. This has freed them from the inflexible neuromuscular system tied to breathing in all quadrupeds. In quadrupeds breathing and running are closely synchronized (one stride per breath). The lungs must be full to support the thorax and absorb the impact of the forelimbs. This opened the way for natural selection of utterances not tied to single breaths, permitting speech, the human laugh (Provine 2004:521), and as argued in this study, singing. Singing, as explained in 2.4.2, requires fine respiratory control and uses all of the respiratory muscles. The size of the thoracic canal is a proxy for the evolution of breathing control (Morley 2003:92; MacLarnon & Hewitt 2004:189; Mithen 2005:227; Fitch 2006a:196) as it reflects the number of nerve cells that control the breath via the intercostal and the abdominal muscles.

`

A bigger thoracic canal indicates that more thoracic spinal nerves were available to control the breath. This larger thoracic vertebrate canal relates in a straightforward way to a larger thoracic spinal cord and an increase in gray matter. Humans have relatively more gray matter, the nerve cell bodies that form the core of the spinal cord, than other species. The white matter, where nerve cell bodies are connected and which comprises a much larger cross-sectional area, of humans is similar to other species. Thus the larger thoracic canal and inferred increase in gray matter mean that local innervation provided by the thoracic spinal nerves increased, and presumably the size of the thoracic spinal nerves (MacLarnon & Hewitt 2004: 189).

MacLarnon & Hewitt (2004) compared the vertebral canal size of fossils of 3 Australopithecines, *Homo ergaster* (Nariokatome boy), 4 Neanderthals and 2 early modern humans with those of 38 species of modern primates. *Homo ergaster's* vertebrae differ from those of modern humans in that the spines were longer, and the canal that enclosed the spinal cord was narrower. These details imply that fewer nerve cells served the chest muscles and diaphragm (MacLarnon 1993, MacLarnon & Hewitt 2004). MacLarnon and Hewitt (2004) satisfactorily address the concerns of Frayer & Nicolay (2001) that the vertebral dimensions of the *Homo ergaster* are actually modern, and that thoracical vertebra do not innervate breathing (see also Morley 203:92-93). They also eliminate other possible reasons for the enlargement of the thoracical canal. For example prolonged running does not require complex neurological control. The enlarged thoracical canal relates to the evolution of breathing control. *Homo erectus* and *Homo ergaster* did not have modern type control over breathing (MacLarnon & Hewitt 2004:191).

MacLarnon & Hewitt (2004) found that only the thoracic region of Neanderthals, early and contemporary modern humans showed a deviation from the pattern of primates. The thoracic vertebral canal remained at primate size during earlier human evolution, and expanded some time after 1.6 million years ago. It is suggested that the common ancestor of Neanderthals and modern humans, at around 600 000 years ago, evolved human-style breathing control. Morley and Mithen do not discuss this finding in relation to speech and not singing. However this adaptation may have been crucial to enable singing around 600 000 years ago.

3.6.3   The evolution of the middle and inner ear

The advent of computed tomography (CT)-based analyses has opened up new avenues of research on the auditory capacities of hominins (Stokstad 2003). There are a limited number of studies on the

`

middle and inner ear bones and its significance for the origins of music. Mithen (2005:129) mentions the work of Moggi-Cecchi on evolution of the bones of the middle ear. These bones amplify and convey sound to the inner ear, as explained in Chapter 2. A part of the stapes (one of the three ossicles of the middle ear), known as the footplate correlates to the range of audible frequencies that can be heard (Moggi-Cecchi & Collard 2002). The larger the footplate, the better detection of low frequency sounds (:262). The footplates of the stapes bone of Stw 151, an australopithecine from Sterkfontein, and *Homo habilis* are ape-like, perhaps indicating an enhanced ability to detect higher frequencies compared to modern humans (Moggi-Cecchi & Collard 2002). Martinez *et al* (2004:9976) notes that this study should be considered with caution because it is based on a simplified model that relies on only a few anatomical variables. Martinez *et al* (2004) undertook several measurements of the middle ear ossicles and concluded that the auditory capacities of *Homo heidelbergensis,* dating to 350 000 years ago, from Sima de los Huesos, Spain, were modern. It thus appears that members belonging to *Homo Heidelbergensis were* able to hear similar frequencies than anatomically modern humans.

The evolution of the vestibular system of the inner ear is of particular significance in the evolutionary origins of music. Morley (2003: 96-100) provides an extensive discussion of the evolution of the vestibular system of the inner ear. He concludes, "the first significant developments of auditory anatomy occur with *Homo ergaster*, 1.5-1.7 Mya, and also seem to be related to a shift to a fully upright posture." Mithen (2005:144) discusses inner ear in passing and also concludes that the vestibular system has evolved a modern configuration at 1.6million years ago with the development of bipedalism. This capsule has an apelike position in Australopithecines, in *Homo habilis* it is monkey-like, but in *Homo ergaster* a modern configuration in which an almost 90 degree rotation has occurred is found (Spoor *et al* 1994). This indicates that complex movements such as running and jumping could have been performed by *Homo ergaster*. In 2.4.3 it has been noted that the changes in vestibular system of the inner ear may have been instrumental in the perception and production of rhythmic movement. It is important here to note that the change in the vestibular system was in place by 1.6million years ago, and if a link between rhythmical movement and the vestibular system can be confirmed, this would mean that rhythmical movement could have been possible at this time.

`

3.6.4 Habitual bipedalism, dancing and running

There is a limited degree of discussion on the implications of bipedalism for dancing in Mithen and Morley's work. As discussed in 2.2.4, the aability to dance is related to habitual bipedalism. Partial bipedalism was in place as early as 6 million years ago (Richmond & Jungers 2008). Partially bipedal hominins could walk upright, but retained powerful arms and curved fingers for arboreal movement. Habitual bipedalism evolved around 1.6million years ago. A remarkably complete *Homo ergaster* individual (KNM-WT 15 000 from Nariokotome, Kenya) (Wood 1992) dating to around 1.6 million years ago already has modern limb proportions. This indicates that *Homo ergaster* was fully or habitually bipedal (Richmond *et al* 2002). The central cores or bodies of the lumbar vertebrae of *Homo ergaster* were broad and circular, with flat, weight bearing surfaces that enhanced balance and strength in the vertebral column. A visible waist between the rib cage and pelvis shows that movements in the upper part of the body were fully independent from those in the lower body (Frayer & Nicolay 2000). The independence of the upper and lower bodies freed the body up for spontaneous movement and the synchronization of rhythmic body movements (Trevarthen 1999).

Mithen (2005:153) discusses Bramble & Lieberman's (2004) finding that the bipedal adaptations of *Homo ergaster* may also have evolved in relation to endurance running. Carrier (1984) and Trinkhaus (1984) have already in 1984 made this argument, but it seems to have attracted limited attention in relation to the origins of music. The long, spring-like tendons like the Achilles tendon, longitudinal arches of the feet, compact feet and small toes, relatively long legs, the type of knee joints and well developed butt muscles, the gluteus maximus are all essential for running but not for walking. This, according to Mithen, would have expanded the potential for gesture and body language. A case can be made that long distance running was much more important for the evolution of musical capabilities than simply expanding the potential for gesture and body language, as is further discussed in Chapter 6, Section 6.5.2. It may have been the foundation of the rhythmical control of movement and the voluntary control of the breath that typify musical activities, but this is an argument that needs to be fully developed in an extensive study. It is however, possible to surmise that dancing movements would have been possible by 1.6million years ago.

`

3.7 DISCUSSION

Two recent multidisciplinary accounts of the evolution of a musical protolanguage have been reviewed with the aim of extracting data relevant to the evolutionary origins of music. Mithen and Morley both see the roots of musilanguage in primate behaviour. It has been concluded here that the research on primate musical abilities has not been extensive enough to draw firm conclusions on its relationship to the origins of music. Primate grunts, "song", bimanual drumming and rhythmic body movements hold considerate promise for finding the precursors to hominin musical behaviours. It was further found that none of the proposed structural properties of musilanguage provides a firm base for the further investigation into the evolution of musical capabilities. In addition, it has been maintained that the evolutionary rationale most relevant to the development of music, is the social bonding hypothesis, as is discussed in more detail in Chapter 6.

The fossil evidence for the evolution of the biological capabilities to produce and perceive musically is sparse. The descent of the larynx plays an important role in the scenarios of Mithen and Morley. However, the research of Nishimura shows that the descent of the larynx involved two stages, the first of which was shared by chimpanzees and must antedate 6 million years ago. It is not known when the second episode of descent took place because none of the indicators used by Morley and Mithen, e.g. the flexion of the basicranium and hyoid morphology reliably relate to the position of the larynx. There are thus no grounds to infer that the descent of the larynx has been tied to the evolution of bipedalism or for proposing a co-evolutionary relationship between the bipedalism, the descent of the larynx and learning to control new types of motor sequences related to speech (or singing). It is not possible to infer on the basis of the fossil record when hominins were able to produce a larger range of sounds made possible by a lowered larynx and an extended vocal tract.

Habitual bipedalism cannot be related to the lowering of the larynx, but it was a crucial event that made musical expression possible. The capability for habitual bipedal locomotion evolved around 1.6million years ago. This freed the upper and lower bodies for more extensive movement, but more significantly, it allowed running that, in turn could have led to dancing. Habitual bipedalism also caused changes in the vestibular system that regulates and coordinates body movements. This system has been implicated in the rhythmical processing abilities that underlie human musical activity. Habitual bipedality further influenced the breathing system and has set in train the development of extensive control of the muscles involved in respiration. The fossil evidence shows

`

that the thoraxical canal enlarged some time between 1.6million and 600 000 years ago. This means that the neural connections to facilitate modern voluntary breathing patterns evolved within this window of time.

Mithen and Morley discuss the lowering of the larynx and breath control in the context of evolution of speech and not singing. Yet the lower larynx that made a larger range of sounds possible, in association with increasing breathing control may have been instrumental in the evolution of singing, not speech. Singing utilises the full extent of the vocal tract (Vaneechoutte & Skoyles 1998; Skoyles 2000). Singing also requires more intensive respiratory control than talking because long sequences of different tones and articulations must be made on a single out breath (Skoyles 2000). *Homo ergaster* represents a watershed in terms of pre-adaptations and adaptations for musicality. In Chapter 2 modern music has been defined as an intentional act that involves complex learned vocalizations combined with whole body movements synchronised to a beat. The evolution of bipedality and the vestibular system after 1.6million years ago would have made it possible to dance, or to execute whole body movements synchronised to a beat. This does not constitute modern music because dancing has to be combined with musical vocalizations. There is almost no fossil evidence that can be related to the capability to vocalise musically. It can be said that it was possible to control respiration in modern ways around 600 000 years ago, and to hypothesise that this was related to singing. More evidence however is necessary to confirm this hypothesis. It is concluded here that the fossil evidence suggests that protomusic consisted of rhythmical whole body movements, perhaps, as argued in Chapter 2, combined with simple vocalisations such as grunts. Chapter 4 investigates whether there is any evidence from the brain sciences to confirm this hypothesis.

`

CHAPTER 4: THE EVOLUTION OF MUSIC AND THE BRAIN

4.1 INTRODUCTION

In Chapter 3 it was found that the period between 1.6million and 600 000 years ago was crucial for the evolution of the capabilities that underlie rhythmical movement and singing. A question that is further investigated in this Chapter is whether there is any information from the brain sciences that supports this finding. A further aim of this chapter is to investigate and evaluate the information on the evolution of the brain and musical processing as presented by Mithen and Morley. Their accounts of brain evolution emphasize different aspects. Morley's view on the development of neural paths for musical processing involves the evolution of separate, lateralized capabilities for emotion and vocal expression on the one hand, and motoric and rhythmic expression on the other hand. He (Morley 2003:221; 2002:208) concludes that between 1.75 million and 300 000 years ago increased vocal control was accompanied by lateralization of prosody and emotional functions in the right hemisphere. According to him, the neurological pathways to control laryngeal and orofacial muscles, specifically involving the nucleus ambiguus, are related to left hemisphere mechanisms that are thought to have first emerged with *Homo habilis*. These left hemisphere developments are tied to motor sequences associated with rhythmic behaviour that eventually evolved to support linguistic syntactical functions. Here the enlargement of Broca's area, argued to be related to the control of motor sequences, instigated the left brain processes. Mithen's discussion of neural relationships underlying musilanguage is limited (Fitch 2005b:488) and he emphasizes that we still have much to learn. Yet his view of the musical protolanguage as being underpinned by relatively discrete modules in the brain is dominant throughout (Mithen 2006:99) and the eventual integration between these modules forms the basis of his explanation for the evolution of grammatical language.

In this chapter, a discussion of the methods that are used to understand current musical and prosodic functioning in the brain is followed by an analysis of the proposals of Mithen and Morley on the neurology underlying pitch, rhythm and emotion in prosody and music. Even though the interest in this study is in the origins of music, their proposals on prosody are included to investigate the commonality between music and prosody from a neural point of view. Their evolutionary paths for

`

brain evolution and musical processing are further evaluated in the context of additional relevant literature.

4.2 METHODS USED TO UNDERSTAND BRAIN FUNCTIONING

Neuropsychologists and cognitive neuroscientists study brain abnormalities to infer how the brain operates normally. Patterns of impaired and spared performance for example reveal the structure and organization of musical and linguistic functions (Rapp 2001). The identification of neurophysiological dissociations plays an important role in this approach. If two functions (e.g. pitch versus rhythmical processing) are selectively disrupted, it is inferred that separate neural systems are involved. In the following example from Peretz *et al* (1998:136) it is inferred from dissociation that "…if a patient can no longer recognize the musical input but is able to recognize its emotional tone, then the dissociation pattern argues for the processing autonomy of the recognition and emotion systems. The inference is even stronger if a different patient is found to display the reverse dissociation; that is a loss of emotional appraisal in the presence of intact recognition. ..[this] constitutes instances of double dissociation, which in turn provides strong support for the independent pathway proposal". Crawford *et al* (2003:357) note that the concepts of dissociation and double dissociation "continue to be defined very loosely" and that more attention needs to be paid to the methodology behind their identification.

The clinical pathologies that have been used to infer the brain correlates of musilanguage include amusia or the loss of musical ability (Mithen 2006:99), auditory agnosia or disruption of the perception of speech, animal noises, music and other sounds (Morley 2003:111), aphasia or the loss of the ability to speak intelligibly (Morley 2003:106) and aprosodias or syndromic deficits in comprehension or production of affective prosody, or both (Monnot *et al* 2004). The methods used to study brain functioning include electroencephalography (EEG), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). EEG reading is the only brain imaging method that directly taps neuronal activities, but it cannot visualize sub-cortical systems. Methods like PET and fMRI measure variations of cerebral blood flow in three dimensions in the brain of a subject performing a particular task (Peretz 2001) and can therefore visualize sub-cortical functioning. The spatial resolution of neuroimaging techniques has improved with the development of methods like diffusion tensor imaging (DT) (Preuss 2006) and there are new methods for investigating neural functioning in infants (Paterson *et al* 2006). The methods used to study brain structures and function in infants include dense array electroencephalography/event-related

`

potentials (EEG/ERPs), near infra-red spectroscopy (NIRS) and an emerging technique, arterial spin labelling (ASL) (Paterson *et al* 2006:1088).

These technologies provide an indication of brain functioning, but doubts have been raised on how to measure neural activity of complex functions, and on the assumptions made in computing results (Sidtis 2000). Peretz & Zatorre (2005) advocate an integration of neuropsychological and brain imaging results, but it is crucial to be aware that similarity of activation does not necessarily mean that the underlying operation is similar or that the recruited brain areas are critical to performance in a given task (Peretz & Zatorre 2005:91). There are sometimes discrepancies in data from brain lesions studies and brain imaging studies. Patel (2003) has noticed that whilst lesion studies show that music and language capacities can be partially dissociated, brain-imaging studies show that they share the same neural networks. The discussion in this Section demonstrates that caution is advised when inferences on brain functionality is made on the basis of neuropsychological and brain imaging results. Even though Morley (2002:203) acknowledges the constraints of the neuropsychological and neuroimaging methods, his evolutionary path for the brain is almost solely based on these studies. In the 4.3 it is examined whether the processing of pitch, rhythm and emotion in music and prosody occur in localised fashion, as suggested by Morley and Mithen.

4.3 THE LOCALISED PROCESSING OF PITCH, RHYTHM AND EMOTION IN MUSIC AND PROSODY

4.3.1 The processing of pitch

Morley (2002:206) notes that there is a tendency for tasks relating to pitch and pitch discrimination to be right-hemisphere dominated. It is indeed one of the better-established facts in the neuropsychology of music that the right temporal neocortex plays a particularly important role in the computation of pitch relations (Peretz & Zatorre 2005; Peretz 2006). Both brain lesion studies (Peretz 2006; Peretz & Zatorre 2005) and neuroimaging studies (Zatorre *et al* 2002; Peretz & Zatorre 2005) show that the right temporal neocortex is involved in the computation of pitch relations as they change over time. The right hemisphere however, is not solely involved in analysing pitch relations. When listeners analyse pitch relations using contour representation to discriminate melodies, the right temporal gyrus plays a critical role, but when contour cues are not available and interval information is required, both the right and left temporal structures are involved (Peretz & Zatorre 2005:92). This cooperation between the hemispheres in the integration

`

of contour with interval information in melodies is also observed with magnetoencephalography (MEG) in neurologically intact people (Patel & Balaban 2000) and infants (Balaban *et al* 1998). There is therefore no simple relationship between the right hemisphere and pitch processing, but there could be an "overall right-sided asymmetry" (Peretz 2006:20) for pitch processing.

Peretz' arguments are also used by Mithen, but to propose that the development of musical abilities depends on the evolution of musical modules. Peretz (2006) proposes that musical modules (neural networks) responsible for the fine-grained processing of pitch exist. For example, neural networks that specialize in the processing of scale structure in melodies occur but the locality of the networks is unknown (Peretz & Zatorre 2005:93). Much of her research is aimed at demonstrating that a musical module exists separately from a speech production module. She supports her hypothesis by several case studies that demonstrate a double dissociation between speech production and lyrical production. Whether this double dissociation relates to an evolved module and whether it is relevant to the evolutionary origins of music, is further discussed in 4.4.2.2.

There are few studies available to evaluate Morley's (2003:221) suggestion that "the production and perception of tonal prosodic sounds of vocalisation were (and are) handled predominantly by the right hemisphere …". Juslin (2001:321) for example discusses cases that show that the right hemisphere of the brain processes voice prosody, but it is not clear which of these were designed to study pitch in particular. Van Lancker & Sidtis (1992) show that the right hemisphere is *more* (but not solely) involved with pitch processing in prosody than the left hemisphere. Patel *et al* (1998) and Nicholson *et al* (2003) confirm this finding. From this discussion it can be seen that both the left and the right hemispheres are involved in pitch related tasks in music (and prosody), but that there is an overall tendency for the right hemisphere to dominate. Whether the neural processing of musical melody and speech prosody is similar, is not well established. There is no firm basis to regard the ability to process pitch in music (and prosody) as an evolved right brain function.

4.3.2 Rhythm and brain processing

Morley (2003:109) does not discuss the processing of musical rhythm to the same extent as that of musical pitch, but he equates the neural computation of musical rhythm with that of syntactic sequencing in the left brain. This suggestion can be questioned, as the segmentation of an ongoing sequence into temporal groups and the extraction of beat in music are probably the result of distinct processing components, as neuropsychological and neuroimaging data show (Peretz & Zatorre

`

2005:94) - meter is supported by the right hemisphere and grouping by the left hemisphere. Mithen (2005:62) does not link the processing of musical rhythm to lateralised function, but he notes that a group of modules in the musical system are concerned with temporal content (rhythm and metre). The significance of modules in music evolution is discussed in 4.4.2.2.

Both Morley (2003:217) and Mithen (2005: 142) consider rhythmically entrained corporeal movement as an integral part of protomusical expression. Morley relates rhythmic bodily movement to left brain processing (2003:217, see also Levitin & Menon 2003). He (Morley 2003:113, 173, 180) further relates the evolution of rhythmic movement abilities to the Intrinsic Motive Pulse and Intrinsic Motor Formation of Trevarthen (1998; 1999). Trevarthen suggests that there is a human "Intrinsic Motive Pulse" (IMP) comprising of the brain stem, basal ganglia and the limbic structures of the emotional motor system. An Intrinsic Motor Formation (IMF) system is responsible for integrating functions of attention, learning and physiological actions of expression, including the synchronisation and co-ordination of sequences of movements in newborns and adults. Unfortunately the physiological and functional parameters of the IMP and IMF are unclear, as Morley (2003:113) remarks and therefore its implications for the origins of music could not be further investigated. Nevertheless, the basal ganglia seem to be important in the distributed circuit involved in estimating temporal intervals and the coordination of patterned movements (Patel 2006:101), but the basal ganglia handles interval timing and motor control functions in many species (Patel 2006:101), not only in humans. This discussion illustrates that bodily rhythmic processing is not a simple process and that much research must still be undertaken to understand its neural functioning.

However, as emphasised in Chapter 2 (2.4.3) the link between rhythm and movement is crucial for understanding the origins of music, a theme that is further developed in the following chapters. There is a strong motor component to the mental representation of musical rhythm (Peretz & Zatorre 2005:95). The cerebellum plays and important role in facilitating rhythmical bodily movement, as Mithen, and not Morley, observes. Mithen (2005:152) refers to studies by Parsons and team (Parsons *et al* 2003) who found that musical rhythm is processed in a distributed way in the brain, and that it activates the cerebellum. The cerebellum (the evolutionary oldest part of the brain, or the 'reptilian brain') may hold crucial clues to the rhythmical expression of bodily movements. Levitin (2006:170-171) explains in more detail why the cerebellum is important in rhythmical movements – it is involved in tracking the beat, musical preference and emotion. It is

`

meaningful that there are direct projections (bypassing the auditory cortex) from the inner ear to the cerebellum (Levitin 2006:180). These projections play a role in the movements that are made to orientate the body in relation to sound. As will be discussed in 4.4, the cerebellum, together with the prefrontal cortex, has evolved to be larger in humans than in other primates (Rilling 2006). It is unlikely that rhythmical processing can simply be related to the left brain as Morley suggests – it involves several brain areas simultaneously. The relationship between rhythmical movement, the cerebellum and the basal ganglia needs to be further explored in future studies on the origins of music.

Prosodic rhythm is not well integrated in the evolutionary paths proposed by Morley and Mithen. Although rhythm is cited as an important element of prosody, rhythm is only discussed in the musical sense. A cursory literature review to determine whether a certain hemisphere dominates in prosodic rhythmic processing gave conflicting results. Van Lancker & Sidtis (1992) found that prosodic temporal information in prosody is extracted by the left brain. In the studies of Patel *et al* (1998) and Nicholson *et al* (2003) it was found that prosodic rhythm is a right brain function. The rhythmic task was to observe the difference between the pauses of clauses (timing shift pairs). The placement of the pause changed the meaning of the sentence. For example, the meaning of "Henry, the little one eats a lot" differs from "Henry the little one, eats a lot". This example is given to show that prosodic rhythm in this sense is different from rhythmic pulse in music.

There is no justification for Morley's (2003:217) suggestion that the left hemisphere is concerned with "the production and processing of rhythm and meter in speech and of rhythm and timing in music". The discussion in 4.3.1 and 4.3.2 supports the conclusion drawn in Chapter 3 (3.3) - that music and prosody have little in common in terms of structure. There is too much ambiguity in the neural processing of pitch and rhythm in prosody and music to support their common neural origin in musilanguage.

4.3.3 Emotion

Mithen (2005) discusses music and emotion extensively, but he does not relate it to brain functioning, except for noting that pitch and temporal modules filter their output through an emotional expression analysis module (:63). In this he follows the research of Peretz and colleagues (Peretz 2001; Peretz & Coltheart 2003; Peretz & Gagnon 1999) who relate the recognition and experience of emotion in music to the functioning of a separate emotional component or module.

`

The issue of modularity is addressed in 4.4. This section aims to determine whether the hypothetical affective nature of musical expression could have been the result of the evolution of mainly right hemisphere functions and the sub-cortical emotional limbic system as Morley (2002:204; 2003:217) proposes.

4.3.3.1 Lateralised emotion processing

Morley mentions that emotional sounds are processed bilaterally, but that there is preponderance for emotional processing to be a right brain function (Morley 2003:217). According to the right hemisphere theory (Demaree *et al* 2005) the right hemisphere is dominant for all emotional processing regardless of valence or pleasantness (Schwartz *et al* 1975; Parr & Hopkins 2000). This idea has found support in a study undertaken by Blood *et al* (1999). Blood *et al* (1999) found that the evaluation of valence activated neural structures located primarily in the right hemisphere.

The valence theory on the contrary posits that emotion is not simply processed by the right hemisphere. This theory suggests that positive emotion is predominantly associated with left hemisphere activity whilst negative emotion is associated with predominantly right hemisphere activity (Davidson & Erlichman 1980; Parr & Hopkins 2000). Studies undertaken to identify the cortical specialisation relating to happy and sad judgements in music (Peretz 2001; Schmidt & Trainor 2001) found that changes in both tempo and mode purported to indicate happiness resulted in greater relative left frontal activations, whereas changes in both tempo and mode in the sadder direction resulted in greater relative right frontal activation. There are also ear-asymmetries in listeners that indicate that the left hemisphere contributes more to the perception of positive emotions and the right hemisphere to negative emotions. McFarland & Kennison (1989 cited in Panksepp & Bernatsky 2002) reported that happy music played to the left ear (thus preferentially stimulating the right hemisphere) tends to increase body temperature, while negatively valenced music had the opposite effect. It should be noted that the valence theory holds for chimpanzees as well. Parr and Hopkins (2000) found that chimpanzees react in the same way as humans to emotional stimuli (although the stimuli did not use musical sounds) – negative emotions resulted in a higher tympanic membrane temperature (Tty) in the right hemisphere. Tty is an indirect site from which to measure brain temperature and it is strongly influenced by autonomic and behavioural activity. It is thus unlikely that the hemispheric lateralization demonstrated when evaluating valence in musical tasks has any evolutionary significance.

`

Most of the studies that Morley (2002) cites for the right brain processing of musical emotion discuss prosodic emotion in speech. The dominant hypothesis for prosody processing is that the right hemisphere dominates for the production and perception of emotion (Ross 2000). Morley discusses the research of Snow (2000) that has shown that the right brain is involved in emotional analysis of speech prosody. Neuroimaging studies confirm that activity in the right frontal regions increases in prosody perception tasks (references in Gandour *et al* 2003:150). The experiments undertaken by Mitchell *et al* (2003) and Buchanan *et al* (2000) found the same trend. Mitchell *et al* (2003) demonstrate that emotional prosody in terms of happy and sad is processed by areas of the right frontal hemisphere, especially lateral temporal lobes. In this study subjects were required to discriminate between sentences describing happy and sad scenarios. Mitchell *et al* (2003) mentions that it is important to separate lexical information and emotion in sentences to make sure that it is the emotion that is conveyed that is judged. In their experiment, a phonetician recorded the sentences and controlled pitch direction, height and range. Thereafter the semantic information was removed and the sentences were largely unintelligible, creating 'pure emotional prosody stimuli' (: 1412). This example shows how musical emotion and prosodic emotion differ in experiments. While Adolphs (2002: 172) corroborates that the right hemisphere is disproportionately important for perceiving and recognizing emotional prosody, he also notes that substantial bilateral processing is involved as well, because the recognition of emotion from the voice draws on multiple prosodic cues processed by systems that are neuroanatomically separated.

The case for right brain processing of prosodic speech seems secure, but both hemispheres are involved in the processing of emotion in music. The studies on emotional recognition in music mentioned here combine many types of stimuli or cues, making it difficult to pinpoint the brain areas involved in emotion processing. Panksepp & Bernatzky (2002) note that there are many problems in determining how happy and sad music affects the brain, not least how to control whether the measured effect in fact represents a direct link between the stimulus and response. Moreover, Zatorre *et al* (2002:45) note that functional specialization emerges at multiple levels and interact in complex ways. There is thus no basis for Morley's argument that the right hemisphere is responsible for prosodic and music emotional processing (Morley 2003:217).

4.3.3.2 The limbic system and musical emotions

The idea that musical emotions are processed by the sub-cortical limbic system (Morley 2002:204) is well established. There are a variety of studies that shows that music stimulates deep sub-cortical

`

regions of the brain (Peretz *et al* 1998; Blood *et al* 1999; Blood and Zatorre 2001). However the certainty that music is processed by the limbic system diminishes if it is taken into account that different researchers define the limbic system differently (Toates 2001). There is still no consensus on which areas of the brain belong to the limbic system (Le Doux 2002). A sensible way to regard the limbic system is to appreciate that these structures do play a role in emotion, but they do so in interaction with structures outside the system, and the structures of the limbic system play roles in processes other than the emotional (Toates 2001:333).

Musical emotional experience is not a simple matter of limbic processing because emotional circuits are widely distributed in the brain. Panksepp & Bernatzky (2002:137) explains that "the emotional circuits, and hence the resulting neural resonances, are widely distributed in the brain, resembling a tree-like structure, with roots andtrunk-lines in subcortical areas, and branches interacting with wide canopies in cortical regions". The Blood and Zatorre (2001) study also mentions an increase in blood flow in areas of the so-called limbic system *and* frontal cortex when "chills" are experienced in association with moving music. These areas include the dorsal midbrain, ventral striatum (which contains the nucleus accumbens), amygdala, insula, and orbitofrontal cortex. Another example of multiple brain areas involved in processing music is a PET and CBF study by Blood *et al* (1999). The PET readings and regional cerebral blood flow (CBF) changes show that several distinct paralimbic and neocortical regions of the brain are activated when the valence (pleasant/unpleasant) of dissonance is evaluated. These studies (Blood *et al* 1999, Blood & Zatorre 2001) were undertaken on musicians.

Menon & Levitin (2005) recorded the "emotional" brain processing of non-musicians in reaction to classical music. They used fMRI and functional and effective connectivity analysis to study how the different brain regions react. The brain regions became activated in a particular order: the auditory cortex, then the frontal regions followed by the mesolimbic system. The sub-cortical regions that became involved include the nucleus accumbens, the ventral tegmentum area and the hypothalamus. There is a "tight coupling" between these areas during music processing. The cerebellum and basal ganglia were active throughout. The nucleus accumbens plays an important role in pleasure and addiction – it is the center of the brain's reward system (Levitin 2006:185). It is further involved in processing monetary rewards, drugs and humour. This study was the first to show direct involvement of the hypothalamus in the processing of music. "The tight ways in which the reward

`

and affective systems appear to interact with the cognitive and autonomic systems in response to music listening are a key contribution of this study." (Levitin & Menon 2005:182).

The inferior colliculus may also play a key role in the affection we have for certain sounds and music. Panksepp & Bernatzky (2002) argue that because the inferior colliculus, mandatory 'way-station' for auditory processes, mediates affective processes and because this region is rich in opioid receptors (in the rat brain) the inferior colliculus may mediate attachments we develop to certain sounds and types of music. The inferior collicus and the nucleus accumbens may thus be of critical importance in the evolutionary origins of music because it is rich in opioid receptors and because it is involved in the production of dopamine, important in positive mood regulation, as will be discussed in Chapter 6.

Neither Mithen nor Morley mentions the importance of the nucleus accumbens and the inferior colliculus and its relation to neurochemicals for the origins of music. The discussion in this section has shown that there are no clear trends in how 'emotion' (disregarding for the moment the ambiguity of the term) is processed in prosody and music. Morley's hypothesis for musical and prosodic emotional processing and its evolution cannot be accepted.

4.4 EVALUATING MITHEN'S AND MORLEY'S BRAIN EVOLUTIONARY PATHS

In Morley's proposed evolutionary path the ability to process music evolves as localised lateralized functional areas in the right and left brain. The right brain and limbic functions are thought to be evolutionarily more ancient than the left brain functions (Morley 2003: 107). Mithen (2005:274) similarly proposes that the modules that process pitch and emotion are evolutionary older than other musical modules. Morley suggests that modern vocal and perceptual versatility and complexity can be attributed to *Homo Heidelbergensis*, 400 000 – 300 000 years ago, while Mithen suggests that by 50 000 years ago the independent modules that have developed to support musilanguage integrate to support music and fully grammatical language.

The following discussion evaluates these proposals for brain evolution against the background of broader evolutionary theory. Firstly, it is explored (4.4.1.1 - 4.4.1.4) how the human brain differs from other primate brains and whether there are any special evolved parts of the brain that may be relevant to musical capabilities. In 4.4.2 it is discussed whether lateralization and modularisation are evolved functions, or whether it is part of the brain's direct response to contemporary stimuli.

`

4.4.1 Unique elements of *Homo sapiens* brains

4.4.1.1 Larger brains

Human brains are larger than those of apes. Modern brain size was achieved by 200 000 years ago (Groves 1989; Davidson 1999) when it reached the modern range of variation around 1350 cc. There are different opinions on whether humans developed their large brains gradually or as a punctuated event. Morley (2003:169) like Deacon (1997:344; 2003) sees a gradual, incremental increase in relative brain size in the evolution of *Homo*. Others describe the increase as a punctuated event (Groves 1989; Aiello 1996b). This view is favoured here. Brain sizes remained stable between 1.6million and 600 000 years ago in terms of brain size relative to body mass (Ruff *et al* 1997; Conroy *et al* 2000). Mithen (2005:158) also notes that there was one million years of effective stasis in brain expansion after *Homo ergaster* first appeared. The most significant increase in encephalisation (the ratio of brain to body size or brain to body mass) took place after 600 000 years ago. Middle Pleistocene fossils which date to between 600 000 and 300 000 years ago such as Bodo, Kabwe, Petralona, Arago and Dali have absolute and relative brain sizes larger than those of *Homo erectus* (and *Homo ergaster*) (Bräuer *et al* 1997; Rightmire 1998).

Humans developed a disproportionally large brain because the genes that code for brain growth changed through evolution. It caused the brain to grow as if it belongs to a primate of a much larger body size. Experimental studies undertaken by Deacon (1997; 2000) show that prenatal brain growth rates in humans, macaques, cats and pigs proceed at essentially the same rate, but body growth rates for macaques and humans are significantly below that for cats and pigs (Deacon 2000:286). Primate encephalisation is therefore the result of a reduction in postcranial growth processes, and not an increase in brain growth. But in humans, another change has taken place and the brain continues to grow as if it belongs to a larger primate species (Deacon 1997). This expansion of the brain is largely confined to the structures on the dorsal half of the embryonic brain – the cortex therefore expanded out of proportion and this has consequences for cortical organization and cortical-sub-cortical relationships as will be discussed below. This bigger brain does not indicate higher cognitive abilities (Roth & Dicke 2005) such as musical abilities. From an evolutionary point of view, an increase in brain size is significant because it indicates reorganisation (Deacon 1997; Weaver 2005).

`

4.4.1.2 The subdivisions of the human brain

Natural selection has modified the patterning of the human brain to deviate from the rules of primates in general (Rilling 2006:74). The fundamental structural divisions of the brain and cerebral cortical regions are controlled by regulatory genes (Edelman 1992). Recently molecular biologists studied how human genes differ from those of other primates and found that humans differ from chimpanzees not so much by differences in the makeup of genes, but, as suggested by the regulatory hypothesis, in how active those genes are.

Data from comparative studies using brain-imaging techniques, pathological lesions and experimental surgery show that humans have an overall larger neocortex, specifically the prefrontal cortex, and cerebellum (Rilling 2006:69). The neocortex is part of the cerebral cortex (the uppermost layer of the cerebral hemispheres) (Rilling 2006:71). The major patterns of cortical wrinkles or gyri can be used as landmarks to identify various regions of the neocortex e.g. the frontal, parietal, temporal and occipital lobes. It is often contended that human frontal lobes are disproportionately large (Rilling 2006), but they are in fact smaller than expected for a primate of our brain size. It is the prefrontal cortex, a part of the frontal lobes, which is proportionally larger. There is increased branching of the dentrites in the prefrontal cortex in humans. Rilling (2006) explains that neurons receive input over their dentritic surface and therefore this branching complexity in the human prefrontal cortex probably reflects integration of a larger number of inputs and increased intra-cortical connectivity. The increased connectivity is reflected in the larger white matter of the human prefrontal cortex, whereas the gray matter volume is not different. These regions are involved in those cognitive functions that humans excel in, for example symbolic thinking, working memory, social behaviour and language. The temporal association cortices are also proportionally larger than those of great apes and this is also reflected in the white matter. The temporal lobe is involved in the perception of sounds and processing of tonality, pulse and dissonance (Peretz & Zatorre 2005) and in the processing of speech (Paterson *et al* 2006). The cerebellum is part of the hindbrain (Mithen 2005:292) and is intimately connected to the cerebral cortex (Rilling 2006: 69). Like the neocortex, the cerebellum consists of cortical gray matter overlying white matter. The cerebellum is involved in motor operations and cognition (Rilling 2006: 65). The larger human prefrontal cortex and cerebellum may have a relationship to musical capabilities, but this relationship is not clear and is bound to be complex and multifaceted.

`

The enlargement of Broca's area in *Homo habilis* is considered important for the increase in manual dexterity in *Homo habilis* and the eventual development fine vocal control (Morley 2003:119). A *Homo habilis* (KNM-ER 1470 from East Turkana) endocast (casts taken from the inside of fossil skulls) shows that Broca's area is enlarged. There are several reasons why inferences on the prominence of certain areas on endocasts should be treated with caution. A more prominent area does not necessarily indicate an increase of that area, but could simply reflect an overall enlargement of the brain (Gannon and Laitman 1993; Gannon *et al* 1998). It should also be kept in mind that Broca's area accommodates more than one distinct processing network (Marcus *et al* 2003; Peretz & Zatorre 2005:15). Morley's proposal that the evolution of Broca's area is important in the development of fine vocal control and therefore musical expression can therefore not be accepted.

4.4.1.3 Species-specific connectivity

Neuroscientific research undertaken from the 1970's to the present has little to offer on which evolutionary specializations occur in the human brain (Preuss 2006:2) and few details are known. Because the human cortex expanded out of proportion, the cortical-sub-cortical connectivity has changed.   Deacon (1997) notes that one likely effect of this shift in proportions is that it changed the number of motor projections descending from the cortical to sub-cortical region. This may be one of the reasons why all primates have direct cortical motor projections for neurons that control the facial, jaw and tongue muscles. Humans have finer and more control over these muscles than primates (Deacon 2003: 283). While both humans and primates have control of the larynx (Deacon 1992), humans are unique in that they have direct connections that enable more extensive voluntary control of the laryngeal muscles. This is evidence of prolonged intense selection favoring increased vocal abilities in our ancestors. Achieving articulate laryngeal and orofacial control was probably a "million-year project at the very least" (Deacon 2000:283). The descent of the larynx already took place before the split between chimpanzees and humans (Nishimura 2006), and there are no indicators for when the second descent took place. It is probable that the second descent took place for reasons unrelated to the control of vocal production, but that vocal communication took advantage of this change. It is uncertain when the evolutionary processes that influenced increased control of the larynx and orofacial muscles were initiated, but if Deacon is correct, it must have been initiated at least a million years ago.

`

Morley relates the ability to finely control the muscles of the face and mouth and to manipulate the body rhythmically to the evolution of the left hemisphere. However, as discussed in Chapter 2, the control of the large body muscles for rhythmic movement has already developed with the evolution of bipedalism and lateralization need not be brought into play for fine orofacial muscle control, as related above - the disproportionate enlargement of cortical parts of the prefrontal cortex and the concomitant availability of neurons to create more connections between the cortical and sub-cortical regions, sufficiently explain the ability to finely control the orofacial muscles.

It seems then that the code that determines the growth of neurons to connect the cortical and sub-cortical regions is prespecified, but as will be discussed in 4.4.2, the dominance of the left or right hemisphere in certain musical functions is part of the generalised way in which all vertebrates process information.

4.4.1.4 Musical genes?

It is unlikely that a 'musical' gene exists that could have been responsible for the evolution of musical capabilities or a musilanguage. However, it intriguing that "the mutated FOXP2 may compromise a common underlying cause for the speech and rhythm deficits, such as a deficit in sequential temporal processing" (Peretz 2006:15). The much-discussed FOXP2 gene is situated on a small segment of chromosome 7 and seems to play a causal role in the normal functionality underlying language and speech and rhythmical processing (Lai *et al* 2001, 2003; Enard *et al* 2002). All the family members with this mutated gene are impaired in rhythm production and perception and oral movements, but they perform normally on pitch-based tasks (Alcock *et al* 2000; Peretz 2006). Fisher (2006:285-290) extensively reviews the research on the FOXP2 gene and cautions that much more research needs to be undertaken before any conclusive inference can be made. It is also interesting that there are several genes involved in hearing (Clark *et al* 2003). Especially the EYA1 gene is thought to be involved in the development of the outer and middle ear (Abdelhak *et al* 1997), but as for the FOXP2 gene, caution is advised because extrapolating from genes to the phenotype is not straightforward (Preuss 2006).

`

4.4.2 The lateralization and modularization of functionality

4.4.2.1 Lateralization

That lateralized functionality is an adaptation that has been inherited as part of the genetic code, as Morley seems to suggest, accords less well with evolutionary theory. Lateralization is not 'hard-wired' into the genetic code. Lateralization in any function, be it linguistic processing or emotional processing, is probably a consequence and not a cause of a particular function (Deacon 1997). Most of the patterning in brain functioning is a dynamic functional consequence of processing demands imposed by the development of that function. Thus lateralization is more an adaptation of the brain *to* a function than an adaptation of the brain *for* a function. Biases for lateralized functionalities such as processing of prosodic features of speech are driven by the particular function during childhood. These biases are subject to a competitive process that may involve both synaptic competition and competition for establishing memory traces of experiences and behaviours (Deacon 1997:310-311). The reason why for example phonemic functions are processed by the left hemisphere while prosody is processed by the right is because there is competition for the same brain structures (Deacon 1997:314).

This logic is confirmed in a review on lateralization undertaken by Valortigara *et al* (1999). They note that it is known that many species of vertebrates share brain lateralization with humans (see also Toga & Thompson 2003). The need for an asymmetrical brain is widespread among vertebrates and the precise direction of lateralization is determined by epigenetic, motivational and emotional factors (Vallortigara *et al* 1999:167). Brains undertake complex computation problems in a lateralized way because functionality improves if they are segregated into sub-problems that differ in typical ways. Invariances and variances, for example are mutually incompatible problems, best handled by functionally separate systems (:168). Chicks and mammals, including humans, seem to use the left hemisphere for cues that allow stimuli to be assigned to categories whereas the right hemisphere is used for spatial analysis and in response to novelty (Vallortigara *et al* 1999:167-168). That these commonalities are common to birds and mammals indicates that lateralization emerged early in vertebrate evolution. The view that brain lateralization is a feature of most vertebras is not universally accepted – it is contested by researchers like Annett (1985), Corballis (1998) and Crow (2004) who regard lateralization in humans as extreme and species specific. A recent review by Toga & Thompson (2003:43) however notes that lateralization is influenced by pre-and postnatal

`

non-genetic factors and that this complicates arguments for a strictly Mendelian inheritance of laterality.

4.4.2.2 Localisation, modularisation and domain specificity

Mithen (2005:99) primarily relies on the research of the neuroscientist Peretz to conclude that both language and music are constituted by "a series of relatively discrete modules within the brain, some of which are most likely shared, and some of which are dedicated to one of these alone". He (:30) suggests these modules are a reflection of up to two hundred thousand years of independent evolution. From child development patterns he infers that the neural networks for language are built upon or replicate those for music (:70). However, he also accepts that the brain "displays immense plasticity as it grows" (Mithen 2005:84, 257). In a sense, Mithen sits on the fence, as Avorgbedor (2008:30) notices. Yet, his conception of evolution of music in Neanderthals is identical to the reasoning used in his 1996 publication "The evolution of the mind" which argues from an evolutionary psychology point of view (Cosmides & Tooby 1994). One of the tenets of evolutionary psychology is that the association areas of the neocortex contain genetically predetermined modules that generate specific types of psychological strategies. Mithen (2006:100) describes Neanderthals as having a domain-specific mentality that imposed major constraints on their creativity and symbolic thought. The transition from Hmmmmm to modern music and compositional language would have involved the transition from a domain-specific to cognitively-fluid mentality. This allowed the production of musical instruments and elaboration of hominin musical abilities.

That the evolution of modules is genetically predetermined is seriously questioned from a biological point of view. Panksepp & Panksepp (2000:111) go as far as stating that evolutionary scenarios based on evolutionary psychology tread close to neurologically implausible views of the human mind. All mammalian brains contain intrinsic, special-purpose neurodynamic functions (e.g. basic motivational and emotional systems), but there is no evidence for genetically dictated adaptations that "produce socio–emotional cognitive strategies within the circuitry of the human neocortex" (:111). Others researchers that caution against the view that genetic adaptations exist in the neocortex include Deacon (1997), Buller & Hardcastle (2000) and Marshall & Fink (2003). Localization, as Marshall & Fink (2003: S6) observe, is not a fixed property determined by genes: "….functional localization is not such a fixed property of brain regions as either lesion studies or early neuroimaging work might have suggested. Rather, the *neural context* in which an area is

`

active may define the cognitive function". It is thus apparent that predetermined musical modules do not exist in the human brain. Another question that is pertinent in this regard is whether there is any evidence for any other type of predetermined innate musical capabilities in infants? In 4.4.2.3 it is investigated how much is known of the innate capabilities of infants from a brain science point of view.

4.4.2.3 Brain development and innate capabilities in infants

The early developmental processes and the influence of genetic factors on the developmental trajectory of the brain of infants are poorly understood (Paterson *et al* 2006:1088). It is known that the timing of brain maturation of the subregions is specified, but the targets of the neurons not. Gogtay *et al* (2007) have taken brain scans from subjects between 4 – 21 years of age once every two years for a period of 10 years. The results indicate the heterochronicity or differential timing of human cortical development, but the subregions follow temporally distinct trajectories (Gogtay *et al* 2007:8178). This means that the myelination of the subregions follows a specified temporal and spatial pattern (Paterson *et al* 2006:1089). Firstly the evolutionary oldest areas in the base of brain are myelinated and around 8-12 months the frontal, parietal and occipital lobes begin to be myelinated. In the developing cortex the rate of maturation of the white matter reflects increasing myelination of axons that enables more efficient transmission of neural signals (Paterson *et al* 2006:1088). Higher-order sensorimotor regions mature after the lower-order sensorimotor regions and phylogenetically older cortical areas mature earlier than the newer cortical regions.

As the targets of the neurons are not specified, the developmental wiring of brains is influenced by the context in which it occurs. The wiring is 'plastic' (Deacon 2000:286). It appears that the necessary information to wire the brain is recruited from extra genomic sources because the infant is born with developmental mechanisms geared to be open to contextual information and environmental input (Deacon 2000:275). Synaptic wiring is the result of "neural Darwinism" (Edelman 1992, 2003; Calvin 1996) - there is an initial overproduction of neuronal cells, and there is competition for sites to connect to. It is not fully known how the competitive process works to form the final neuronal connections (Deacon 2000:278), but one intriguing aspect is that axons are able to hold on to synaptic connections with a target cell when these fire in synchrony with one another and the target cell. Neural connections are fine tuned by infant behaviour. In the first years of life dewiring, rewiring and myelination take place as the environment influences the child. The infant brain does not function because of innate connections as proposed by the Chomskian nativist

`

tradition (Le Doux 2002:66) because networks of neurons have to be fine-tuned to create and strengthen the connections. These neurons have to be shaped to form circuits and systems. A circuit is a group of neurons that are linked together by synaptic connections whereas a system is a complex circuit that performs a specific function like hearing.

This is in line with the 'epigenetic' conception of the developing organism. Epigenetic is defined as "those modifications to DNA or the nucleosome that do not alter the sequence of nucleotides, but do modify the transcription of genes."(Keverne and Curley 2008: 399). Epigenetic modifications in the brain are non-heritable but they enable long-term stable changes in neural and behavioural phenotypes in response to environmental experiences (: 409). Because input from the environment is so integral to its functioning, brains have been described as bio-environmental or bio-social organs (Gibson 1996; Müller 1996; Greenberg *et al.* 1999). From this discussion it seems that, from a brain science point of view, little is known of innate capabilities of infants. Whether infants possess innate musical capabilities from a behavioural point of view is further investigated in Chapter 5.

4.5 DISCUSSION

It is unsatisfactory that there is so little evidence from brain sciences that can be used to inform on the evolution of musical abilities. More intensive investigation of current brain functioning would not clarify the evolution of dancing and singing abilities because of the sheer enormity of data that needs to be integrated to understand the simplest cognitive task (Fisher 2006). The evolutionary paths for brain evolution proposed by Morley and Mithen are at odds with the evidence from brain science discussed in this chapter. Moreover, the methodology used by psychoneurology and neuroimaging involves assumptions on the relationship between the observed results and actual brain activity and computation of results. This gives reason to treat evolutionary scenarios based on current brain functioning, as proposed by Morley and Mithen, with caution. In addition, much more research is necessary to understand how the brain processes music (Peretz & Zatorre 2005:106).

The musical aspect that is best understood from a neurofunctional perspective is pitch. Neuropsychological and neuroimaging studies demonstrate that most aspects of music and prosody are processed in functionally separate specialised areas of the brain, but these are not strongly lateralized. The literature overview has not confirmed Morley's description of the localisation and lateralization of pitch, rhythm and emotion in prosody and music. The only convincing lateralized

`

function is the processing of pitch relationships in the right hemisphere. The discussion in 4.3 has shown that both hemispheres are involved in the processing of other aspects of musical pitch, musical and prosodical rhythm as well as emotion. The evidence presented in this chapter furthermore reinforces the view, as discussed in Chapters 2 and 3, that prosody and music have little in common in terms of musicality.

Morley acknowledges that the proposed lateralised functioning is not completely clear-cut and that there are exceptions to the rules of lateralization (Morley 2002:206), yet he suggests different evolutionary rates and relationships for these functions. An important conclusion from this chapter is that the genetically pre-specified or pre-determined lateralisation and modularisation that are key elements in Morley's and Mithen's respective scenarios are not part of the genetic code, and could not have evolved as such. Genes determine very few components of the brain and the lateralization and functional specificity of musical functions in living subjects have little relevance for the evolutionary history of those functions. The discussion on brain evolution has highlighted that relatively little is known about species-specific brain specialisations, their genetic underpinnings and their evolution, but neuro-studies show that only the larger sub-divisions of brains are specified by evolution. Individualised functionality is the result of neural Darwinism and epigenetic processes. This plasticity is one of the reasons why Mithen's hypothesis on the evolution of modularity and Morley's hypothesis on the evolution of laterlization are less convincing.

There is no direct evidence from the brain sciences that can be used to support or refute Morley's or Mithen's propositions for the timing of modern musical (and linguistic) abilities. However, the increase in brain size discussed in 4.4.1, coupled to the other biological changes that enabled musical behaviour is meaningful. By 600 000 years ago there was a marked increase in brain size and, as discussed in Chapter 3, the changes in respiration, movement apparatus and ears had been completed by that time. This enlargement would have provided an opportunity for brains to become differently organised and for new cortical-sub-cortical connections to become part of the genetic make-up. A significant aspect of this overview on brain evolution is that the ability to voluntarily control the orofacial muscles is related to the out-of-proportion evolution of the cortex. This enlargement allowed more neurons to connect the cortical to the sub-cortical parts of the brain and it is these connections that are crucial in voluntary control of muscles. It seems that the voluntary control of some muscle groups is at the heart of constructing a mind and body that is able to produce music and speech. The muscles that control breathing and the larynx may have been under

`

voluntary control earlier than the orofacial muscles. Singing requires intensive voluntary control of the breath and of the laryngeal tract.

The brain enlargement within the *Homo* lineage means that brains had to reorganise because more dendrites are necessary to connect the parts of a larger brain (Deacon 1997). The reorganisation of brains in evolution is driven by co-evolutionary processes (Deacon 2003). This means that existing behaviours drive the direction of brain re-organization. It is proposed here that protomusical behaviour was already part of the behavioural repertoire of archaic hominins 600 000 years ago, and that this musical behaviour co-evolved with the brain and aided in the evolution of typically human brains able to think in symbols and to speak. A period of time in which behaviours and the genes co-evolve would have been necessary to fix new changes in the genome. Boone and Smith (1998) and Deacon (1997) suggest that thousands of generations are necessary for fixation.

Morley's scenario for the development of modern musical (and linguistic) abilities is in keeping with this framework. Morley suggests that by 300 000 years ago essentially modern vocal and musical abilities had evolved. Mithen suggests that these modern abilities evolved only by 50-40 000 years ago. The evolution of habitual bipedality at 1.6million years ago was important in the ability to entrain movements to a beat, or to dance. This implies that musical behaviour in the form of rhythmically entrained movement was possible very early in hominin evolution. The ability to voluntarily control respiration implies that perhaps, but less certainly, creatively constructed and manipulated vocalisation (not necessarily melody) was possible by 600 000 years ago.

`

CHAPTER 5: THE MUSICAL BEHAVIOUR OF INFANTS

5.1 INTRODUCTION

Mithen and Morley both discuss the significance of Infant-directed communication in terms of the evolutionary origins of music. Mithen (2005:74) implies that the musicality in Infant-directed communication was part of Hmmmmm and that its underlying mechanisms once belonged to a musical ability that was used to regulate social relationships and emotional states. For example, for Mithen Infant-directed speech (IDS) may "…have provided a further dimension to 'Hmmmmm' and the evolutionary precursor of the first stages of the IDS used by *Homo sapiens* today" (Mithen 2005:204). Mithen (2005:71) discusses the functionality of Infant-directed communication: it serves to engage and maintain the infant's attention, it modulates arousal and emotion, it acts as communicator of feelings and intention and helps in the acquisition of language. One of the aims of this Chapter is to investigate these propositions by Mithen. Infant-directed communication is described and further investigated in Section 5.2 whilst Sections 5.3.2 and 5.5 analyse the functionality of ID-communication in terms of motivation and language acquisition.

Morley (2003:222) mentions that Infant-directed communication may be a specialized perpetuation of a type of non-linguistic interaction used by all individuals in prehistory. The contention is that the infant's sense of rhythm and melody may have a basis in hereditary factors (Morley 2003:114). The greater part of this Chapter is devoted to probing the proposition that the infant's sense of melody and rhythm may have a basis in hereditary factors. If these capabilities were evolutionarily important, they should form part of the innate abilities of newborn infants. Innateness is inferred when a capability appears "early" and "universally" (Peretz 2001:105), if it has been acquired "easily" (Lamont 2005:361) and when such a behaviour functions at birth "independent of environmental influence and experience" (Drake & Bertrand 2000:18). The hypothesis, that the capacity to perceive melody and rhythm and to gain affective information from them is innate in newborns is investigated further in this Chapter through a literature review. In Sections 5.3 and 5.4 the literature review discusses foetuses' and newborns' capabilities to perceive pitch and rhythm and their emotional attachment to ID-communication.

`

5.2   INFANT-DIRECTED COMMUNICATION

Infant-directed communication (ID-communication) is the melodic, slow and repetitive way in which caregivers communicate with infants. Morley (2003:313-315) does not describe ID-communication in depth whereas Mithen (2005:69-84) provides a more informative chapter on this subject. Mithen (2005:69) explains that caregivers communicate with infants in very specific ways – a higher overall pitch, wider range of pitch, longer hyperarticulated vowels and pauses, shorter phrases and many repetitions are used. Although Mithen has captured the traits most often listed for ID-communication, they may not accurately portray ID-communication. Trainor *et al* (2000) undertook a detailed comparison of Adult-directed (AD) and Infant-directed (ID) speech and found only two acoustic differences between the two: ID speech is slower in love-comfort situations and higher in pitch than AD speech.

Infant-directed communication involves more than vocalisation - it combines vocalisation with "gestural and kinesic" aspects (Falk 2004:495). Stereotyped, repetitive and exaggerated facial expressions, head movements, and gestures accompany Infant-directed communication or motherese. Caregivers and infants take turns in using this type of communication and they imitate each other. This is why ID communication is described as a "multimedia duet" that incorporates "synchrony of behaviour and alternation" (Dissanayake 2001:86).

The way in which caregivers speak to infants is similar to the way in which they sing to them, but ID singing has received much less attention than ID speech (Trehub 2001). Infant-directed singing is a human universal and shares many characteristics in terms of pitch and tempo with Infant-directed speech (Papoušek *et al* 1991; Trainor *et al* 1997; Trehub *et al* 1993). All caregivers, including fathers and pre-school siblings, sing to infants in a specific way (Trehub 2001). Compared to when they are singing alone, they sing to infants at a slower tempo, a higher pitch with exaggerated rhythm in a more emotionally engaging manner (Trainor *et al* 1997) to comfort them and to get their attention. Lullabies and playsongs are typically used with infants (Trehub 2001). ID singing plays a critical role in caregiving behaviour (de l'Etoile 2006) as infants show more sustained attention in response to maternal singing than to maternal speech (Trehub 2001; Nataka & Trehub 2004). Maternal singing has an effect on the arousal levels (as measured by salivary cortisol levels) of healthy unstressed 6-month-old infants. Singing heightens attention, leading to the rise in cortisol, but further research is necessary to determine the significance of this finding (Shenfield *et al* 2003).

`

Longhi & Karmiloff-Smith (2004) consider Infant-directed song, rather than Infant-directed speech as the basis of Infant-directed communication. Longhi (2003) undertook a longitudinal microdevelopmental study of infant-mother dyads. She analysed the temporal structure of songs that mothers sing to their infants as well as the infants' responses. As in ID speech, infants listen actively to their mothers' songs and participate in the synchronous timing of movements, anticipating the song's segmental units. Babies respond to ID song with synchronised movements (Longhi 2003). This finding is in contrast to Shenfield *et al* (2003:372) who describe maternal singing as essentially a solo performance. Mothers gradually replace the synchronous physical contact with their infants while singing, with contact using toys and other external objects. This brings about a change from dyadic interaction to triadic interaction through song. Trehub (2001:8) interestingly notes that mothers' performances of songs seem to become ritualized because the pitch level and tempo stay virtually identical when repeated over various periods (in contrast to repeated verbal phrases).

That even two-day old infants born from congenitally deaf parents prefer Infant-directed singing may be an indication of inborn capabilities (Masataka 1999:1001). The ability to perceive pitch and melody, temporal sequences and patterns and emotiononal aspects are seen as innate (Trehub 2003; Peretz 2006). These predispositions and their pre-natal origins are discussed below to ascertain whether they are relevant to the origins of music.

5.3 PRE-NATAL INFANTS

5.3.1 The development of hearing and movement in utero

The foetus is exposed to the internal sounds of the mother's body and this includes her heartbeat or pulsating blood vessels, body movements, impact sounds like footfalls, digestion, breathing and vocalizations (Lecanuet 1996). Hearing can be regarded as the dominant prenatal sensory modality (Parncutt *in press* a). Parncutt (*in press* b) mentions that the cochlea and vestibular system of the inner ear starts to function from 18 weeks gestational age. This is confirmed by Jeffery and Spoor's (2004) research that investigates the labyrinthine morphology of 41 post-mortem human foetuses ranging from 9-29 weeks gestation using high-resolution magnetic resonance imaging. They found that the components of the labyrinth, specifically the cochlea and lateral canal, follow distinct growth trajectories, and that the prenatal labyrinth attains the adult size and shape between 17 and 19 weeks gestation.

`

Airborne sound to the foetus is filtered through the amniotic fluid. It is not known whether in the early foetal stages sound travels directly through the external and middle ear fluids or via bone conduction in the foetal head. Prenatal auditory perception and reaction is a complex field of study but it is known that functional hearing is achieved from between 26 and 30 weeks as the neural pathway between the cochlea and brain stem develops (Joseph 2000). Foetal response to sound is not cortical in origin, as the forebrain (foremost part of the brain) and cortex is not developed at birth (Joseph 2000; Trevarthen & Aitken 2001). The foetus processes sound with a pathway between the brainstem (pons in conjunction with the medulla and the midbrain) and the VIII nerve and cochlear nucleus (Joseph 2000; Glendenning 2005). The neurocognitive processing of sound improves with time. Foetuses respond to external sound (Groome *et al* 2000) as well as the mother's voice (DeCasper & Fifer 1980; Lecanuet 1996; Groome *et al* 2000; Kisilevsky *et al* 2003; Al-Quahtani 2005). By 30 weeks gestational age (GA), foetuses respond to brief episodes of relatively loud (110 dB) airborne sounds (Kisilevsky *et al* 2004) and by 36 – 38 weeks the foetus responds to external sounds by reflexive body movements, head turning and an accelleration of heart rate (Joseph 2000).

There has been little research on foetuses' response to music per se. Kisilevsky *et al* (2004) have tested the reaction of foetuses on auditory stimuli (a 5 minute piano recording of Brahms' Lullaby). The mothers wore headphones through which other music was played to mask the mother from the effect of the music. This was done to ensure that the foetuses respond to the Lullaby and not to the mothers' reaction to the Lullaby. A maturation of music perception over the last trimester of pregnancy was observed. Foetuses in all age groups (28 weeks GA to term) showed some heart rate response to the music stimulus. In the younger foetuses music elicited heart rate acceleration (indicating arousal) at the higher intensities while lower intensities elicited a decleration (indicating attention). The term foetuses reacted with a deceleration of heart rate to both high and low intensity sound levels. Near term foetuses have the ability to discriminate and respond differently between musical stimuli depending on its tempo, loudness and pitch. These results have been confirmed by Al-Quahtani (2005) who played Spanish guitar music to foetuses between 37 and 40 weeks of gestational age.

Foetal movement begins at about eight weeks as the foetus moves through the amniotic fluid and extends and flexes its limbs. Most of the foetus' movements are mediated by the brainstem and these functions occur in a rhythmic, diurnal, and/or reflexive fashion (Joseph 2000). Thelen (1981)

`

describes repetitive, cyclic movements around an axis, and flexion and extension movements as rhythmic stereotypes. These movements seem to be regulated by a maturational trajectory for neuromuscular pathways because its onset is regular. The stereotypical movements peak between 24 and 42 weeks of gestational age. Motor control develops in conjunction with the foetus' movements. In this phase proprioception or the sense of the relative position and motion of the parts of the body develops (Parncutt *in press* a). The motor response adapts faster to new stimuli than for example the cardiac response. In habituation experiments it was found that in foetuses, premature and newborn babies the cardiac response is slower to habituate than the motor response. At birth the infant has a repertoire of movement patterns that are the building blocks for its voluntary actions. Multiple subsystems control action in the different body regions. The spontaneous movements of newborns are asymmetric and are more coordinated on the right than the left side of the body. Male babies move differently from female babies. There is a difference in the lower-body movements of males and females that parallel the adult pattern – males tend to be left biased in their lower body movements. This trend is ascribed to the sex steroid hormones that affect the central nervous system in the foetus and uterine function in the mother (Grattan *et al* 1994). The newborn baby is born with the ability to discriminate and recognise sound, and with the motoric ability to move the limbs at will. The newborn's ability to synchronise their movements and vocalisations with the ID-communication of adults rests on these capabilities.

5.3.2 Motivational and emotional states of the foetus

The foetal environment is critical in directing the motivational and attentional predispositions of the infant (Trevarthen 1998). Trevarthen's work incorporates Panksepp's (1998a,b) notion that the evolutionary older parts of the brain (e.g. brain stem and 'limbic system') are important in directing and influencing behaviour in the foetus and infant. A neural understanding of emotional feelings or affect is one of the keys to understand complex brain phenomena (Panksepp 1998a; Damasio 1999). The foetus may have no 'true' feelings because the forebrain, as it has not fully formed yet, cannot mediate through cognitive processing (Joseph 2000). However not all emotion is mediated by cognition. There are basic emotional operating systems that are typified by cognitive/affective/ behavioural patterns linked to certain neural circuits that start forming in utero (Panksepp 1998b) involving the sub-cortical parts of the brain. The foetus' brain already develops by means of self-organising and neural Darwinian processes as discussed in Chapter 4. In the last stages of foetal development excess dendrites, synapses and neurons that are not stimulated are discarded (Joseph

`

2000). Foetal auditory experience shapes the pathways of the brainstem and midbrain and the foetus thus increasingly becomes responsive to those sounds that it is repeatedly exposed to (Joseph 2000). These sounds generate emotional reactions in humans (Panksepp & Bernatzky 2002), and most likely in the foetus.

Mammals use emotions to self-regulate, unlike vertebrates who use automatic processes. Trevarthen (1998; Trevarthen & Aitken 2001) suggests that an emotional motivation system emerges in the embryo brain. This emotional system regulates attentional orientations and coordinates purposeful movements of the body and its parts. This is in accordance with the ecological approach (Gibson & Pick 2000) that regards organisms as motivated to actively exploit the benefits of the environment. This is possible without the involvement of reflective consciousness (Parncutt *in press b*). To exploit the environment effectively, foetuses must anticipate the consequences of their movements. Parncutt (*in press b*) suggests that the foetus learns by "classical conditioning". An organism has "learnt" when it modulates its future response to an environmental stimulus. When two stimuli repeatedly occur together, the foetus associates the one with the other as in the theory of classical conditioning. It has been demonstrated that the rat foetus becomes conditioned after just four pairings of chemosensory stimuli and Lecanuet (1996) discusses successful conditioning experiments with foetuses and sound.

Parncutt (*in press* a & b) suggests that internal patterns of sound become associated with movement and emotion. This is based on Mastropieri & Turkewitz' (1999) research that has demonstrated that the mother's voice, heart, stomach, footfalls and body movements depend in consistent ways on her physical and emotional state. Sound patterns transmit information about the mother's state more quickly than hormones related to emotional states of the mother. The patterns of sound and movement become associated with hormonal changes by classical conditioning. The sound and movement come to predict the emotional correlates. In this way the foetus will, after many repetitions in this pattern, begin to respond emotionally (conditioned response) to changes in sound and movement patterns (conditioned stimulus) that occurs in utero - the hormonal changes will be anticipated by the sound and movement (Parncutt *in press b*:14). This information is relevant because Parncutt suggests that music is "an evolutionary parasite of prenatal auditory, vestibular and proprioceptive development" as infants will be strongly motivated to engage in behaviours that invoke prenatal sounds, movement and emotion. To further investigate this intruiging hypothesis, it

`

needs to be clarified which particular physical and emotional state relates to which particular sound stimulus.

Trevarthen & Aitken (2001: 17) propose that the foetus is "powerfully conditioned, from the outset by how it chooses or plans to move". Two complementary motive states are regulated – those that maintain the organic functions and those seeking experience to build efficient anticipatory cognitive systems. All regulatory actions are constrained by the capacities of the body. Effective regulation depends upon integrating internal physiology and visceral or autonomic states with the moving body. Therefore the infant is born preconditioned to use its bodily actions and perceptive organs to communicate intent. The infant is born with awareness that is specifically receptive to the subjective states in other people – this is intersubjectivity. This ability is described as being "specialized, innate, 'human-environment-expectant' and socially regulatory" (Trevarthen & Aitken 2001:4). Other animals also show this intersubjective regulation, but in simpler ways. The process by which the foetus develops basic emotions and motivational states is vague, but newborn infants are specifically motivated, beyond instinctive behaviours, to communicate their need for social interaction.

This discussion has shown that it is plausible to regard prenatal experiences as important in shaping the development of the foetus' brain and movement patterns and as playing a role predisposing the newborn to prefer and react to musical communication. Prenatal perception can only have an effect on postnatal survival if the infant has memory of the prenatal sound patterns and activities, which is the case. Infants have transnatal longterm memory lasting for at least five weeks (Lecanuet 1996; Kisilevsky *et al* 2004). Communicating with the infant in motherese prolongs and modifies prenatal memory traces. The newborn infant already has basic discriminatory abilities that, as will be discussed below, aid in eliciting and enabling musical responses. As the auditory system is more mature at birth than the visual system (Lewkowicz 2003) vocal expressions are likely to play a major role in early emotional development (Trainor *et al* 2000: 194).

5.4 THE NEWBORN AND INFANT AND MUSICAL PREDISPOSITIONS

5.4.1 Pitch discrimination and manipulation

Infants have well developed pitch discrimination abilities as they can detect pitch changes of a semitone or less (Trehub 2003:669). However, it cannot be said with confidence that this ability is innate in human infants as apes display similar abilities. Apes can, for example, discriminate

`

between complex tones (Bendor & Wang 2006). However, there is a possibility that the ability to attend to melodic contour (Fishman *et al* 2001; Trehub 2003:669) is innate in human infants. Human infants further prefer sequences of consonant intervals above sequences of dissonant intervals (Schellenberg & Trehub 1996; Zentner & Kagan 1996, 1998; Trainor & Heinmiller 1998; Trainor *et al* 2002; Trehub 2001, 2003). Their perception of consonant (e.g. perfect fifth or fourth) intervals (two simultaneous or sequential tones) is also more precise than that for dissonant intervals (e.g. tritone) (Trehub 2003). Neurophysiological (Fishman *et al* 2001; Tramo *et al* 2001) and behavioural (Zentner & Kagan 1998; Izumi 2000) studies have shown that rodents, birds, monkeys and human infants share the ability to discriminate musical chords in terms of sensory or psychoacoustic consonance. It is not easy to determine whether nonhuman animals share the preference of human infants for consonant intervals. McDermott & Hauser (2004) undertook experiments with tamarins and they argued that the human subjects preferred the consonant intervals while the tamarins showed no preference for consonant or dissonant intervals, leading them to conclude that the preference for consonance is inherent to humans (McDermott & Hauser 2007). However, the method that they followed, as Lamont (2005:360) notes, is suitable to show aversion, but not to capture motivation or engagement in the monkeys or humans. Also, the intervals presented to the tamarins and humans in this experiment were synthesized complex tones with ten harmonics and the bass note of the intervals was middle C. It is not clear whether this is comparable to sensory consonance. The study of consonance and dissonance preference and whether it is innate to humans needs more work and carefully designed experiments of what it is that is measured – how do you measure preference in monkeys if their 'theory of mind' and social preferences are not known?

Human infants have the ability to adapt their vocalisations to those of their caregivers - mothers and infants imitate each other's pitch and match it precisely (Papousek M. 1996). Even premature newborns have the ability to imitate vocally (Trevarthen & Aitken 2001:7). Babies furthermore often respond in "gliding-type" sounds. Infants' imitative vocalizations are frequently musically logical, in particular during song sung by the other (Malloch 1999). As far as could be determined, non-human infants do not make use of imitative and matching vocalisations. Masataka (2007:36) also notes that vocal matching occurs exclusively in adult New world monkeys and great apes, not in infants or juveniles.

`

5.4.2 Rhythmic abilities

Infants share basic rhythmic discriminatory abilities with animals. Infants use temporal information to discriminate their native language - 4 day old and 2 month old infants are able to successfully discriminate their native languages from foreign languages. Rats and tamarin monkeys are also able to discriminate language from different classes when speech is played forwards, but not when it is played backwards (Vouloumanos & Werker 2007). Newborns have not yet been exposed to the cultural preferences of their temporal environment and therefore process tempo and rhythm more 'openly' than 12 month-olds (Stewart & Walsch 2005). Hannon & Trehub (2005a) have demonstrated this in their experiments that show that 6- and 7-month old infants have not yet developed the culture specific preference for certain rhythms that is present in 12-month old infants and Western adults (Hannon & Trehub 2005b).

Infants already display the universal rhythmic processing abilities presented by Drake & Bertand (2001) (as discussed in Chapter 2). Six-to eight-month infants are able to group temporal events in the same way as adults. Infants are also able to perceive a small change in the tempo of an isochronous sequence at two months of age (Baruch & Drake 1997). The preference of humans to naturally perceive and prefer temporal intervals in binary (duple meter) ratios is also present in infants and this seems to be a natural inclination (Bergeson & Trehub 2006). Infants move their heads, arms, chest, and legs stereotypically in a binary way and early bipedal kicking and sucking have binary aspects (Pouthas 1996; Thelen 1981).

The research of Pouthas (1996) noted infants' ability to adapt the rate of spontaneous sucking to that of an auditory sequence. The optimal processing zone and sensitivity to change is highest if events occur every 600 ms (1 ms is $1/10^{th}$ of a second), the tempo of the infants' endogenous clock (Bègue *et al* 2006). Baruch & Drake (1997) have shown that infants react to novelty only for sequences at 600 ms IOI (see also Bobin-Bègue *et al* 2006). Very young infants therefore very effectively perceive and process context-linked temporal stimuli if these are faster than their endogenous tempo (600 ms) and they are predisposed to attend to oscillations with brief periods (McAuley *et al* 2006). Infants of 2 months old can adapt their sucking rate to accelerations of 15% if the reference tempo is 600 ms, but not if the change in tempo is slower than 600 ms. Although newborns and 2-month-old infants can perceive a change in tempo, they cannot adapt their sucking rate to the tempo (Bègue *et al* 2006). However, they do perceive the slower tempo because their sucking becomes more regular (like that of newborns) when it is presented to them. The ability to

`

adapt sucking rate to a slower tempo develops only at 2 months of age after a period of fairly intensive training.

From an evolutionary point of view, it is essential to know whether the uniquely human propensity to search for temporal regularities at a particular rate and adapt their motor behaviour to it, or sensorimotor synchronisation, is already developed birth. The study of sensorimotor synchronisation is a relatively new and limited field of research. Most studies on synchronisation discuss tapping tasks by adults (Repp 2005). There is very little data on synchronisation in infants (Pouthas 1996). Infants have the ability to modify their motor behaviour according to changing tempos. Malloch (1999) and Trevarthen (1999) discuss the capability of infants to co-ordinate body movements rhythmically with vocalisations. Rhythmic coordination is the basis for infants' (and premature infants') ability to exactly synchronise with adult's gestures and utterances (Malloch 1999: 36-37). Mothers accompany their singing by simultaneously shaking their heads, rocking their bodies and patting their infant's bodies in a temporally organized way (Longhi 2003). This type of synchronisation is different from the adult's musical sensorimotor synchronisation where actions are planned and executed according to an anticipated beat. However, as will be discussed below in Section 5.4.3 newly born infants are already capable of "interactional synchrony" and the rhythmical matching of actions described here may be one of the developmental bases of musical synchrony and entrainment.

There is a reciprocal relationship between movement and rhythm. Movement influences the way in which rhythm patterns are encoded in 7-month-old infants (Phillips-Silver & Trainor 2005). This was determined by first bouncing infants on either a second or third beat of an ambiguous (unaccentuated) rhythmical pattern. When accentuated rhythmical patterns were presented to them, they preferred those patterns they were taught to bounce to. As discussed in Chapter 2 and 3, the vestibular canals influence the way in which movements are made and monitored (Spoor *et al* 2007). This system develops early (Clark *et al* 1977). Infants take pleasure in vestibular stimulation when bounced or rocked to songs or in play. In the Clark *et al* (1977) experiments, infants between 3 and 13 months were exposed to vestibular stimulation by holding them in an upright sitting position on the lap whilst spinning the chair clockwise and counter-clockwise. The infants' head would be held in different positions to position the semicircular canals in various positions. The gross motor abilities of all of the children improved significantly with this stimulation (Clark *et al* 1977). This interaction between the vestibular system and the auditory system for temporal

`

grouping "is critical for the development of musical behavior" (Phillip-Silver & Trainor 2005: 1430).

The ability to entrain and synchronise movements according to an anticipated beat does not seem to be present in newborns, probably because their capability to control their muscles are immature. Adults optimize the tonal and rhythmical capabilities of newborn infants to achieve "companionable communication" (Malloch 1999:48) that can be described as musical. Systematic interaction between mother and infant is created through rhythm and tone to allow them to express themselves in sympathy with each other (Malloch 1999) as will be discussed below.

5.4.3. Emotional and motivational predisposition for infant directed communication

The intense interest of infants in music is remarkable (Trehub 2003:669; Peretz 2006:17). Most authors regard the emotional preference for Infant-directed communication as innate (e.g. Papoušek, M. 1996; Papaeliou & Trevarthan 1998; Malloch 1999; Trevarthen 1998, 1999; Trevarthen & Aitken 2001; Trainor *et al* 2000; Trehub 2003; De l'Etoile 2006; Peretz 2006). Infant-directed communication serves to regulate and modulate emotional bonding. That the emotional messages in Infant-directed communication are of extreme importance is illustrated by the fact that mothers who suffer from depression have trouble communicating to their babies and babies have trouble in responding to their depressed mothers' speech sounds (Robb 2000; Gratier 2000) and singing (Trevarthen 1998). In the first six months of infants' lives they are capable of vocally expressing discrete emotions (Papaeliou *et al* 2002) and to differentiate prototypical vocal expressions in Infant-directed speech and respond adequately to their categorical messages (Papoušek *et al* 1990). After 6 months they have developed the ability not only to communicate emotions, but also to infer dispositions and intentions and conveying communicative functions to facilitate goal orientated activities (Papaeliou *et al*. 2002).

Newborns and very young infants are highly motivated to elicit attention by adjusting their actions and vocalisations to be similar to those of the caregivers. Interactions in a secure environment with newborns are marked by sustained mutual attention and rhythmic synchrony of short utterances and touching. Infants are socially gifted in integrating perceptual information from many modalities to serve motive states. As early as 2 months infants and caregivers regulate each others' interests and feelings in intricate, rhythmic patterns. Infants' hands make expressive movements in rhythmic coordination with another person's speech, even if the baby has been blind from birth (Trevarthen

`

& Aitken 2001). Miall & Dissanayke (2003:352) describe this capability as "….the coordinating of behavioral-emotional states with another's through temporally organized sequences." The term that they use to describe the ability to create and maintain interpersonal coordination is mutuality (Miall & Dissanyake 2003:352). The importance of "intersubjectivity" and "mutuality" for the evolutionary origins of music is further discussed in Chapter 6.

Newborns (including premature newborns) have the ability to imitate accurately. They can imitate "large tongue protrusions, exaggerated opening of the mouth or eyes, looking back over the head, holding up a hand, extension of one or two fingers, single vowels imitated in a rhythmic burst" (Trevarthen & Aitken 2001:8). This ability is not unique to humans as chimpanzee newborns also have the ability to imitate (Myowa-Yamakoshi *et al* 2004). However, the ability to imitate a broad range of whole-body actions or facial expressions seems to be unique to human infants. The ability to consciously and intentionally represent events through re-enactment through mimicry and imitation is reffered to as mimesis (Donald 1991). It is beyond the scope of this study to evaluate Donald's cognitive theory of the evolution of the mind, but the concept of mimesis is a central part of Mithen and Morley's thesis.

Imitation in newborn infants serves interpersonal functions. These are innovative behaviours and the infant uses concentrated effort to make it happen. It is not spontaneous affective regulative expressions like smiles and crying. In Trevarthen's (2005: 94) words: "Within hours of birth, babies imitate face expressions, hand gestures, shifts of the head and looking or closing of the eyes and simple vocal sounds. In experimental demonstrations, imitating, defined as reproduction of the same form of act as the act presented, is a rather puzzling activity, elicited by a ritual of exaggerated 'modelling' behaviour of an adult interrupted by waiting for a reaction from the infant. Observed in more spontaneous or intuitive encounters by adults who want to communicate with the infant, the infant's responses exhibit the two cardinal features of conscious, self-motivated or intentional behaviour – inter-modal sensory equivalence, and motor equivalence for matching gestures with different body parts.

This imitative ability extends to synchrony with speech patterns. Infants exhibit "near-adult competence at interactional synchrony within twenty minutes of birth" (Benzon 2001:26). Interactional synchrony is "the relationship between a listener's body and the speaker's voice……The listener's body movements lag behind the speech atterns by forty-two milliseconds or less…". It is interesting that interactional synchrony (and self-synchrony) reflect the hierarchical

`

nature of speech – larger gestures track phrases while smaller gestures track words or phonemes. This means that infants must have an innate ability to track and entrain to vocalisations of others (whether spoken or sung) in utero. Benzon discuss the work of Condon (e.g.1986) that shows that a delay in entraining with the actions of others (between 100 to 266 milliseconds instead of 42 milliseconds) is related to pathologies like autism and dyslexia in children. This interactional synchrony reflects the operation of the phylogenetically older brain structures, because it is only these structures that are functional at birth. Tightly synchronized actions with others thus play a vital role in the way in which the cerebral cortex matures. This illustrates that a central aspect of musicking, to move in synchrony with another, is present at birth.

The discussion in Section 5.4 has shown that human infants are not born as musical savants. Human infants do, however, have the unique and innate ability to imitate an extended range of whole-body actions, facial expressions and vocal patterns. The infant's capability to achieve the interactional synchrony through imitation is significant for the understanding of the origins of music. This capability, coupled to the extreme motivation of infants to elicit attention by adjusting their actions and vocalisations is at the heart of music making.

5.5 MUSICAL PREDISPOSITIONS AND THE ACQUISITION OF LANGUAGE

Infants use musical characteristics to make sense of speech intonation and to attend to intended meanings. ID-communication's initial role is to communicate the caregiver's feelings and intentions, but as the infant matures, it plays a role in facilitating language acquisition (Papoušek M. 1996; Monnot 1999; Trainor *et al* 2000; Trevarthen & Aitken 2001; Mithen 2005). Infants use their ability to imitate to discover the structure of language. Vocal imitation in the auditory domain is much easier than other forms of imitation because the individual serves as an 'acoustic mirror' to himself - he can hear his own vocal output and compare it with its memory of another individual's vocal output (Fitch 2000:265).

Grunt vocalisations may play an important role in the acquisition of language, but this aspect of infant behaviour has received little attention. Grunts are among the earliest vocalizations expressed in infants (McCune *et al* 1996:28). They are exerted in reaction to movement or effort in relation to adjustment of the vocal tract for protective closure and food ingestion. The increase of grunts at about 3 months of age is associated with the effort of initial upright face-to-face interactions with the mother. It is also expressed with active movement and early object manipulations at about 6

`

months. Grunts decline proportionally with the onset of babbling around 7 months. It is suggested that grunts are instrumental in the acquisition of language (McCune *et al* 1996 and references therein) – grunts first occur autonomically in response to metabolic demand, and then as a communicative vocalization.

Infants' experience of their own autonomic grunt influences the shift to intentional utilization in communication - the attention that the infant gives to the grunt vocalization is crucial. Motor activities, such as modification of respiratory rhythm or maintaining an upright pose in infants are associated with intensive attention, which leads to glottal activity and grunts. A learned association develops between the attentional state and the grunt vocalisation, leading to the production of grunts when the infant experiences a sense of volition, even though the metabolic demand may be absent. These grunts may become associated with internal mental states and environmental events in specified ways, and as such may function as symbols. There is a similar shift from autonomic grunts in vervet monkeys to expressive grunts. Vervet infants produce early grunts associated with effortful movement, but these same grunts are produced in adults when they are stationary, but observe others moving.

Infants play creatively with sounds. In this "vocal play" (Papoušek M. 1996:104) they repeat and modify sounds discovered by chance through their spontaneous vocalisations. Vocal play peaks by the end of six months and develop into babbling (canonical babbling) and spontaneous singing. Infants use babbling that consists of repetitve vocalisations accompanied by rhythmic hand gestures, as discussed below, to discover grammatical language. The ability to babble is not 'hard wired' or automatic. Its development depends on the auditory experience of the infant (Ejiri & Masataka 2001; Masataka 2007). Babbling is first produced around 7 months of age (Petitto *et al* 2004). It consists of well-formed consonant-vowel syllables that develop into the rapid glottal and articulatory movements used to produce grammatical speech. Babies use their sensitivity to the temporal constrasts and rhythmic characteristics of speech to discover language's phonological inventory. A feature of babbling is its rhythmicity (Ejiri & Masataka 2001). McAuley *et al* (2006) discuss evidence that shows that rhythmic patterning and isochrony is the fundamental trait used to track linguistic utterences. The hypothesis, that early language learning depends on an attentional bias for periodic sound patterns (a periodicity bias) (Lewkowicz 2003), is further supported by the finding that infants who cannot discriminate rapid auditory cues are language impaired in later life (Benasich & Tallal 2002).

`

On the other hand, it has been suggested that the movements of the jaw or mandbular oscillations associated with chewing and sucking determine the syllabic structure of babbling (e.g. MacNeilage 1998). MacNeilage's frame/content theory proposes that the physiological properties of the jaw 'frames' or guide the content of syllables in babbling. Therefore there would be a finite number of fixed consonant-vowel sequences. However, there are large individual differences in the consonant-vowel sequences of babbling (Petitto *et al* 2004:46) and this hypothesis is still vigorously debated (Fitch 2000:261; Petitto *et al* 2004:46).

Vocal behaviour is accompanied by rhythmic action of the hands in the period just before babbling begins. In babies that hear normally, this "manual babbling" disappears, but deaf babies use manual babbling exclusively (Petitto and Marentette 1991). Babies that have not been exposed to an environment where langauge is spoken seem to use a qualitatively different manual babbling (Petitto *et al* 2004). Petitto *et al* (2004) tested two groups of babies with normal hearing abilities - one of the groups has never been exposed systematically to spoken language (because the parents are deaf) while the other group has been exposed to language. Babies not exposed to a normal linguistic environment produced linguistic manual babbling at approximately 1 Hz, and non-linguistic manual babbling at a higher frequency of approximately 2.5 Hz (Movement segments were determined by an innovative Optotrak system and calculated in terms of frequency in Hertz). The infants exposed to speech only used manual babbling at 2.5 Hz. The manual babbling of the babies of deaf parents contained linguistic components, but those of the 'normal babies', not.

This indicates that babies require patterned linguistic (not strictly auditory) cues from their environment to develop babbling that forms the basis of linguistic competency (Petitto *et al* 2004). The infants' brain is sensitive to temporal groupings and needs the particular rhythmic input patterns in the structure of language to activate the rudimentary motor programs to produce them. These motor programs are not restricted to a particular modality – speech and manual movements in infants are "equipotential articulators" and it depends on the baby's exposure which modality will be used (Petitto *et al* 2004: 69). Infants use a musical aspect, rhythmicity, to couple vocalisations (initially grunts) with the motor actions of the orofacial muscles, or to develop linguistic sign language.

Grammatical language, whether signed or spoken, is used to create a symbolic web that structures the lives of humans. This symbolic world is created for infants, in conjunction with them (Anderson 2005). The body and brain work together in a dynamic system to achieve this. The infant

self-create the symbolic awareness that comes with grammatical language in an autopoetic process (Trevarthen & Aitken 2001; Trevarthen 2005). Autopoiesis implies that a living system couples to its environment and this triggers structural changes in the system (Thelen & Smith 1998). The types of changes that are possible depend of the nature of the structure, however. Through the process of "emergence" novelty is created that is often qualitatively different from the phenomena out of which it emerged. The symbolic mind is thus an emergent phenomenon dependant on adequate "space" in the neocortex. The development of an individual able to support a human mind (language and modern music) has to take into account physics, biochemistry and biology of the nervous system as well as the nonlinear dynamics of neural networks. Brain structure is epigenetic and developmental processes in the brain are configured to take place in an increasingly organized and systematisized basis (Thelen & Smith 1998). From an evolutionary perspective neither linguistic grammar nor music are predispositions, but the intentional utilization of rhythmicity to discover regularities is.

Babies learn through experimentation and discovery and language and music are created and realized through physical action. The embodiment hypothesis is becoming increasingly prominent in neuroscience, cognitive and develomental science (Clark 1999; Smith 2006). In this paradigm intelligence emerges as a result of the interaction of an individual with the environment and as a result of sensorimotor activity (Smith & Gasson 2005). Thelen's work for example (Smith 2006) shows that there is no "program" for a specific motor behaviour like speech that pre-exists in the infant's genes. In this paradigm linguistic communication is a self-organizing behaviour that develops as a solution to a problem. Through embodied learning the infant discovers subtle linguistic regularities, builds on the progess it has already made and utilizes the input of "smart social partners" to develop the symbols that form the basis of language (Smith & Gasson 2005). In spite of these explanatory paradigms, understanding the development of human symbolic cognition remains a challenge (Clark 1999: 350).

5.6 DISCUSSION

In this Chapter a number of Mithen's and Morley's ideas on Infant-directed communication and the evolutionary origins of music were investigated. Even though Mithen provides a well-grounded description of Infant-directed communication, further reading (Trainor *et al* 2000) has shown that the most salient aspects of Infant-directed speech are a slower tempo and higher pitch in love-comfort situations. The literature review has also shown that, as Mithen (2005:79) mentions,

`

Infant-directed singing may be more critical in caregiving behaviour than Infant-directed speech as infants show more interest in Infant-directed singing. The response of babies to ID speech and song is notable – they listen actively to their mothers' songs and speech and participate with synchronised movements. Emotional regulation of the infant takes place through rhythmic events, imitation and synchronised actions. When the infant is born, he is already motivated to integrate perceptual information from several modalities to serve motive states, a phenomenon termed "intersubjectivity" (Trevarthen 1999) or "mutuality" (Miall & Dissanyake 2003). This capability may be evolutionarily adaptive as psychological disorders result from insufficient interpersonal action and early childhood learning is dependent on intersubjectivity (Trevarthen & Aitken 2001; Trevarthen 2005).

An aim of this Chapter was to test the hypothesis that infants are born with innate musical capabilities, a possibility mentioned by Morley. From the literature reviewed it was learnt that the infant is born with superior hearing abilities and a sensitivity for melody; with a sensitivity for spotting changes in temporal sequences and an ability to imitate vocally and bodily. It seems that infants are not born with the capability to entrain their movements to an external beat, but they are extraordinarily sensitive to changes in tempo and they can adjust their sucking rate to tempos that are similar to or faster than their endogenous tempo. At birth they already display interactional synchrony with the vocalisations and actions of others. In Section 5.3 it was argued that the strong intrinsic motivation to elicit attention and attain intersubjectivity or mutuality by adjusting actions and vocalisation through imitation in synchrony with adult's ID-communication is precocious and of particular importance in the context of the origins of music.

It is furthermore significant that infants use these musical abilities to acquire grammatical language, as Mithen points out, in passing. On further investigation it became evident that it is the ability to vocally imitate, combined with grunt-like vocalisations, which form the basis of the infant's first speech-like communications. The other musical ability that plays a key role in the acquisition of language is the infant's extraordinary sensitivity to rhythm. Infants discover language's phonological inventory through their sensitivity to temporal contrasts and rhythmic characteristic of speech. The musical aspects of rhythmic recognition and movement and vocal imitation thus play a vital role in the acquisition of language. In this view grammatical language is not a genetically predetermined content-rich module that is switched on during the infant's development, as implied by a modular view of the mind. Grammatical language is acquired dynamically through interaction

`

of the body and brain and through the processes of autopoeisis and emergence (Smith & Thelen 1998). A noteworthy point, as discussed in 5.5, is that babies learn by means of experimentation and discovery through physical action as proposed in the embodiment hypothesis (Clark 1999; Smith 2006).

The embodiment hypothesis is relevant to thinking about the evolutionary origins of music: Infants' three musically relevant 'innate' capabilities or predispositions – imitation through whole body synchronous movement and vocalisation; an extraordinary ability to perceive and adapt to rhythmical cues similar to the endogenous clock; and a strong motivation to elicit emotional attention may have interacted through embodied learning in evolution. It is plausible that protomusic, consisting of imitative rhythmical synchronous movement and grunt-like vocalisations played a role in the development of intersubjective awareness. It can be suggested that natural selection acted on the capability to synchronise actions and vocalisations rhythmically and on the motivation to do so.

This does not mean that Infant-directed communication was the source of musicality, as suggested by Dissanayake (2005: 378). The specific modern human expression of Infant-directed communication probably reflects, rather than contains the seeds of, evolved musical mechanisms to attain mutuality or intersubjectivity. It is more plausible that the tendency to mimic extensively and to synchronize vocalizations and movements with others to achieve emotional convergence evolved in a wider context. In Chapter 6 the "emotional" reaction to music, said to be evolutionary important, is investigated from various physiological perspectives.

`

CHAPTER 6: A BIOPSYCHOLOGICAL INVESTIGATION OF MUSIC AND "EMOTION"

6.1 INTRODUCTION

That musical expression and perception are inextricably linked to emotion is one of the core propositions in Mithen's and Morley's publications. The hypothesis is that the function of musical expression in musilanguage was the communication of emotion: Mithen (2005:2) remarks, "If music is about anything, it is about expressing and inducing emotion". Morley (2003:161-162) argues that a fundamental component of evolutionary fitness is the emotive prosodic elements of vocalisation in music and speech. This is a common notion. Peretz (2006:23) for example states that the study of music as an emotional medium may explain why it is so widespread and useful. Mithen (2005) discusses the emotional effect of music in a short Chapter entitled "Music hath charms and can heal – music, emotion, medicine and intelligence". This Chapter is thought provoking and touches on many important issues, for example the need to identify universal musical emotions and how emotion in music has been identified, but it does not discuss any of the aspects in depth. Morley's (2003) discussion the "emotional" significance of musical expression and perception in his Chapter 6 on the "Rationales for the evolution of the Capacities for Music" covers similar ground. This Chapter aims to investigate biological perspectives to the study of musical emotions to complement the views of Mithen and Morley.

It must be noted that the research on "emotion" in music is fraught with difficulties. One of the most problematical aspects is the difficulty of adequately defining "emotion". Much academic effort has gone into defining emotion, but as yet there is no consensus on a definition. Kleinginna and Kleinginna (1981) for example listed 92 different definitions of emotion and highlight the scepticism around its study. The main problem is fuzzy semantics (Cabanac 2002; Kappas 2002; Panksepp & Bernatzky 2002). In 6.1 the approach to the study of emotion in this study is outlined – it is proposed that "emotional" reactions to music include somatic or bodily responses. Whether psychological research of emotion in music has the potential to inform on the evolutionary origins of music, is an issue addressed in 6.2. In Sections 6.3-6.6 various physiological reactions to music, including that of the sympathetic nervous, the biochemical and neuronal firing systems are

`

investigated. Including these aspects considerably broaden the explanatory potential of the "emotional" effect of music. The goal of this Chapter is to investigate whether there are predictable universal "emotional" psychophysiological reactions to musical sound that could have been meaningful in the origins of music.

6.2 THE APPROACH TO MUSICAL "EMOTION" IN THIS STUDY

Darwin's publication of "The Expression of the Emotions in Man and Animals" in 1872 has been the cornerstone of studies on the adaptive role of emotions, as mentioned by Mithen (2005:85). After Darwin and throughout the twentieth century there was very little discussion of emotion in studies in evolutionary studies of the mind and behaviour (Damaiso 1998). This is no surprise, as emotions were considered to be too esoteric, elusive and subjective to study scientifically. This has changed to the degree that a general science of emotion (Kappas 2002) has emerged. However, there is still a lack of appropriate research paradigms to study the emotional effects of music (Juslin & Sloboda 2001; Scherer 2004). In this study the developments in the biology of emotion (Damasio 1996; Le Doux 1996) guide the approach followed to investigate the emotional effect of music and its evolutionary significance.

Le Doux's description of emotion as the process by which the brain determines or computes the value of a stimulus (2002:206) is an example of a biological approach to emotion. He describes emotion as a process that consists firstly of a bodily response, followed by a feeling and an action. The bodily response is associated with changes in physiological parameters. Subsequently, a feeling emerges when awareness develops that something is felt. A feeling is thus the conscious experience of an emotion (Le Doux 2002:225). This often leads to action. The detection (feeling) and reaction processes may take place automatically and independently of conscious awareness. For example, a person takes action when in danger, jumping away from an approaching vehicle, and only afterwards physiological reactions (increased heart rate) and feelings of fear are noticed. This conception of "emotion" as a mental state as well as a somatic response is widely held (e.g Toates 2001; Cabanac 2002). Damasio (1996) similarly describes emotion as an assessment of body states through direct sensing.

Scherer's (2004:240) discussion of musical emotion is in line with this biological conception of emotion. He defines musical emotion in terms of three reaction components that he terms "the emotional response triad" - physiological arousal, motor expression and subjective feeling. Mithen

`

also considers emotions in the biological sense and he comments that some universal emotions are "wired into the human genome by our evolutionary history" (Mithen 2005:86). In the following Section (6.3) it is explained why, in this study, the psychological study of musical emotion is regarded as a limited source of insight into the evolutionary origins of music.

6.3 THE PSYCHOLOGICAL EFFECT OF MUSIC

Morley (2003: 151) discusses the emotional effect of music in terms of Juslin's & Sloboda's (2001) categories of "extrinsic" and "intrinsic" affect. "Extrinsic" affective reactions are elicited as a result of associations of the properties of the music with previous experiences and events. Therefore researchers like De Paula & Dourish (2005) regard this psychological association with music as socially constructed. It is thus the experience of the individual and his associations with the music that elicit a psychological response or a "feeling". These types of experiences have no relevance to the study of the evolutionary origins of music. Juslin & Sloboda further mention the possibility that there are "intrinsic" emotional properties of music. The intrinsic emotional properties of the music "are those which elicit in the listener an emotional response as a direct consequence of structural properties of the music itself" (Morley 2003:151). The question is whether such intrinsic emotional properties in music actually exist.

Cooke (1959) is one of the earlier generation researchers who argued in this fashion that there is a close relationship between structural characteristics of music and certain emotions (Mithen 2005:85; 90-92). Cooke (1959) attempted to "illustrate how particular [tonal] relationships express particular emotions" (in Mithen 2005:91). For example a major third was considered an expression of pleasure while a minor third was regarded as an indication of painful emotions. However, the work of Cooke is based on an outdated notion of "emotion" and does not include current understanding and critical evaluation of the concept. It is important to realize that researchers like Cooke who attempts to identify a relationship between a structural aspect of music and an emotion sometimes uses Western concert music as a medium. The "emotional meaning" of this music would be lost on non-Western societies, just as some 18th century colonial observers were perplexed by Bushman music and declared it to be 'unmusical' (as remarked by Stow 1905). Music may have its own method of coherent emotional expression as Cooke (1959:212) maintains, but only for those who have been inducted into the system.

`

It is often (Peretz 2001; Peretz *et al* 1998; Peretz 2006; Panksepp & Bernatsky 2002) noted that musical emotions are quickly and easily perceived by members of the same culture. Ordinary adult listeners need less than a quarter of a second of music to reliably distinguish the tone of the whole musical excerpt as "happy or sad" and they exhibit a high degree of consistency in their judgements. There are a number of studies in which researchers aim to identify an intrinsic relationship between a musical aspect and an emotion. For example Peretz *et al* (1998:117) maintain that the two structural properties that are known to be relevant to the happy-sad distinction in music are tempo and mode. According to them slow tempi, or few beats per minute, and minor modes tend to evoke sad moods and fast tempi and major modes happy moods. Dalla Bella *et al* (2001) and Peretz (2006) also report that six-year-old children and adults have the ability to use tempo and mode (minor/major) to distinguish between happy or sad. There are inconsistencies in these results as in the Dalla Bella study 3-4 year olds could not distinguish between happy and sad melodies, but Peretz (2006:23) reports that children of this age can recognise happiness in art music of their culture. Peretz *et al* (1998) further conducted an experiment with a non-musician woman (I.R.) and four normal controls, to test whether they could recognize emotion in music. IR has lost the ability to perceive 'cognitive' elements of music such as pitch and time processing. In this experiment fast tempo is defined as a quarter note value that varies from 80-255 Metronome Marking, (M.M.) and slow tempo as music between 20 and 100 M.M. It was found that the subjects could reliably distinguish between happy and sad music based on tempo and mode and they did so within as little as half a second from the beginning of the music. Human expressiveness and instrumentation were found to have negligible effects on the evaluation of music as happy or sad. Although the subjects could distinguish between happy and sad automatically and effortlessly, the way in which they accomplished these judgements are not understood well (Peretz *et al* 1998:133).

The properties of music discussed above may have intrinsic emotional effect in the sense that there is a demonstrable link between an aspect like tempo and an emotion like sadness. This does not mean that the emotional reaction is predetermined or consistent. The studies are few and differ in design - the types of music, age of subjects and background of subjects differ from experiment to experiment. The experiments are conducted on small groups or individuals, usually from a Western background.   Therefore the link between an emotion (e.g.happy/sad) and musical property (fast/slow) is bound to be culture specific. Another aspect that precludes drawing an evolutionary meaningful relationship between a musical aspect and emotion is that measuring emotion is complex. Convnetionally the reportable subjective experience that occurs during an emotional

`

state, for example the feeling of fear, anger, happiness or joy is monitored by means of standard psychological measures like Anxiety Inventories or Self-Assessment tools to assess valence or pleasantness. These tests are cultural constructs in themselves. It can only be concluded that there is as yet no empirically observable universal intrinsic link between a structural musical aspect and an emotion. In the next section it is investigated whether the relationship between music and physiological indicators show more promise to understand the evolutionary origins of music.

6.4 THE REACTION OF THE SYMPATHETIC NERVOUS SYSTEM TO MUSIC

Physiological measurements of "emotional" responses to music include assessments of changes in heart and pulse rate, electrodermal or galvanic skin response, respiration rate, blood pressure, muscular tension and tone, blood volume, skin temperature, gastric mobility, pupillary reflex and blood oxygen (Bartlett 1996:345). These physiological responses occur mostly without voluntary control or awareness. The hypothesis in most of the studies that investigate the effect of music is that certain elements of music have a regular and predictable influence on mind and body states. Below it is investigated whether there is an observable universal relationship between a number of physiological indicators and an emotion.

Bartlett (1996) reviews 62 studies undertaken between 1906 and 1995 that consider the relationship between sound (the type of music is not always specified) and heart rate. It appears that heart rate does not automatically covary with music, but asking subjects to experience the emotion expressed in the music with the eyes closed, resulted in a much stronger relationship between heart rate and music (Gomez & Danuser 2004). It should be commented here that such studies could only be of significance for understanding the evolution of the origins of music if the experimental design is sufficiently rigorous to isolate the effect of music.

One aspect that is of interest in the context of this study is that certain specific rhythms result in entrainment of the heart rate (Abrams 2001). This type of synchronisation takes place "because 2 similarly vibrating rhythms interact and resonate at the same frequency" (Aragon 2002:54). Aldrige (1993:17) reports a study by Bason in which it was found that an audible click played at a precise time in the cardiac cycle, had the ability to increase or decrease the heart rate up to 12% over a period of time of up to 3 minutes, but when the click was not within the time range of the cardiac cycle, there was no entraintment.

`

Although music has an effect on respiration rate, there is no clear pattern. Sometimes stimulative music causes an increase in respiration rate, but soothing classical pieces and subdued jazz also cause an increase in respiration rate (Bartlett 1996). Musical training has an effect on whether rhythm influences respiration rate (Aldridge 1993). Musically trained subjects experience significantly greater coordination between breathing and musical rhythm than nonmusically trained subjects. Most experiments investigate respiration rate in combination with other physiological responses. For example, Gomez and Danuser (2004) investigated the relationship between arousing and calming music, environmental sounds and skin conductance, heart rate, and respiratory patterns. They found that more rapid breathing, increase in mean inspiratory flow and minute ventilation is associated with music that result in arousal (Gomez & Danuser 2004). In the main the relationship between respiration and music is not clear. This pattern of inconclusive results holds for the relationship between musical sounds, blood pressure and blood volume. The relationship between musical sound, muscular tension, motor activity, skin temperature and gastric activity has also been tested. The results from studies in all these categories are inconclusive (Bartlett 1996). There is no simple relationship between musical sound and these physiological reactions.

One type of physiological reaction to music often discussed is the "chills", originally described by Goldstein (1980). Whether the "chills" represents a universal response of the automic system's to musical sound (Rickard 2004; Craig 2005; Grewe *et al* 2007) has been investigated through the skin conductive response (SCR), or galvanic skin response (GSR) (Bartlett 1996; Khalfa *et al* 2002). It was established that chills do not consistently elicit changes in GSR or SCR and that it is not an automatic reflex-like reactive pattern to certain musical features (Grewe et al 2007:301). What Sloboda (1991) and Pankesepp (2005) respectively call "sad pieces" aroused the chills most effectively while Grewe *et al* (2007) found, in contrast, that pieces perceived as pleasant elicited the chills. Their (Grewe *et al* 2007:312) main conclusion, and one that is important for understanding the effect of music on the human psyche and body, is that it is not the musical features, but the attention given to and associations with the music that are important in arousing the chills. Davies *et al* (2003) similarly note that the personal and subjective associations with music, more than the structural elements of the music, play an important role in the for example, anxiolatic effect of music.

It is furthermore interesting that reported emotions sometimes contradict physiological reactions to music. Subjects may describe music as highly relaxing, soothing and sedative, but their

`

physiological reactions may indicate the opposite effect (Aldrige 1993:21; Bartlett 1996). It is again probably the memories associated with the music that elicits the reported emotion. A specific physiological response may be associated with more than one affective state. For example, increased heart rate may be associated with fear as well as happiness. Meyer described this dilemma as early as 1956 and referred to it as undifferentiated responses. There seems no reliable relationship between music, the sympathetic nervous system and emotion that can be used to research the origins of music.

6.5 THE BIOCHEMICAL REACTION TO MUSIC

The psychological and the physiological studies discussed so far have limited potential to understand the emotional impact of music. This section is aimed at determining whether the biochemical reaction to music could provide information more relevant to the evolutionary origins of music. The relationship between emotion, biochemistry and music has received increasing attention in the past ten years, thanks to the development of psychoneuroimmunology (PNI). PNI considers the interactions between behaviour, the brain, the endocrine system and the immune system (Ader 2000; Viljoen *et al* 2003). There is a two-way communication between the immune system and neuroendocrine system that starts in vitro (Viljoen *et al* 2003). PNI is also referred to as psychoneuroendocrinoimmunology (PNEI) because the neural and immunological systems have receptors in common. Immune stimulation influences the central nervous system (CNS) while neural activity and cerebral lesioning affect the immune response (Mašek *et al* 2000), but many processual details are still unknown (Haas & Schauenstein 1997). The influence of music on the immune and neuroendocrine systems is discussed below.

6.5.1 Music and immune competence

Most music emotion studies have measured the effect of music on two primary immune markers – cortisol and immunoglobulin A (IgA). Cortisol is a steroid hormone that is produced in the adrenal glands adjacent to the kidneys. Cortisol regulates blood pressure and cardiovascular function and also the use of proteins, carbohydrates and fats in the body. The amount of cortisol is often expressed in relation to DHEA. DHEA is another a steroid hormone, and is related to testosterone and oestrogen. Both cortisol and DHEA exert potent immunological effects. The cortisol to DHEA ratio is regarded as a good indicator of adrenal steroid immunomodulatory function (Bittman *et al* 2001:44). Cortisol secretion increases in response to stress and arousal (Rickard 2004).

`

Shenfield *et al* (2003) investigated the effect of maternal singing on the arousal of healthy, non-stresssed 6-month-old infants. They found that the songs regulated salivary cortisal levels - those with lower baseline levels of cortisol showed "modest"cortisol increases whereas those with higher baseline showed modest reductions. Maternal singing thus has an effect on the arousal levels of healthy unstressed 6-month-old infants. The increased arousal is related to the heightened attention in reaction to singing. Fukui & Yamashita (2003) report that cortisol levels dropped when subjects listened to music. Khalfa *et al* (2003) examined the changes in salivary cortisol levels in response to relaxing music as compared to silence after a psychologically stressful task. Listening stopped the cortisol increase, but in the group who experienced silence after the stressful task, cortisol kept on increasing for 30 minutes. Arousal responses to music are differentiated - sad music for example elicits an increase in arousal in some and a decrease in arousal in others (Krumhansl 1997).

Another way to research the influence of music on the immune system, is to measure the levels of the immune agent found in human saliva, secretory immunoglobulin-A (S-IgA). Salivary immunoglobulin A is a protein that is considered to be the first line of defense against respiratory infections. Playing of music termed "Heart Zones" to hospitalised children resulted in an increase S-IgA after one session (McCraty *et al* 1996). Music listening combined with certain types of imagery is reported to be effective in promoting immunity (Abrams 2001:224). Mc Craty *et al* (1996) found that music especially designed to be soothing as well as self-induced states of appreciation produced a significant increase in automic function (50%) and salivary S-IgA (55%). The two interventions together however resulted in a substantially greater (141%) increase in S-IgA.

Musical activity in a group context has a strong immunological effect. Drumming, in combination with pleasant group activity that emphasises "camaraderie, group acceptance, lighthearted participation and nonjudgmental performance" has resulted in enhanced immunity as measured by increased DHEA-to-cortisol ratios, increased NK cell activity (NK cells are large granular lymphocytes) and increased LAK cell (cytolyticy) activity (Bittman *et al*. 2001:38). It was also found that singing in a choir led to an increase in positive affect and S-IgA in an experiment undertaken by Kreutz *et al* (2004). The activity of singing had a more positive affective and immunonological effect than just listening to music (the subjects were amateur choristers). Beck *et al* (1999) similarly describe that active singing led to strong increases in S-IgA (up to 350% in 25 % of the subjects) in professional choral singers. A decrease of cortisol was experienced during

`

rehearsals, but an increase in cortisol occurred during the more stressful permorfance. Interestingly, there is also a relationship between singing and immunity in other species (Duffy & Ball 2002).

The mechanisms that mediate the immunoenhancing effects of music and positive emotional states on S-IgA and other immunological indicators are unknown (McCraty 1996:172). The main finding of this Section is that there is a strong positive relationship between immune competence and the activity of making music in a group. The immune system can be regarded as a sixth sense organ that provides information on the environment to the brain (Blalock 2002) and as such it exerts a powerful influence on the development and the evolution of behaviour. It is in this sense that the immunological reactions to music are relevant to its origins.

6.5.2 Music, neurotransmitters and modulators

Le Doux (2002: 41) explains that the chemical molecules that are released when neurons are activated are called neurotransmitters. The neurotransmitters drift across liquid filled synaptic spaces, and come into contact with postsynaptic cells where they cause electrical responses. The neurotransmitters function in cooperation with other chemicals like modulators. According to Herlenius & Lagercrantz (2001:30) it is difficult to distinguish between transmitters and modulators but one distinction involves the speed with which they act. Transmitters are fast acting whereas modulators have slower and longer-lasting effects. Peptides, hormones and amines are examples of modulaters (Le Doux 2002:57). Neuropeptides like the opioids, endorphins and enkephalins alter pain sensations and mood. The monoamines include substances like serotonin, dopamine, epinephrine and norepinephrine (Herlenius & Lagercrantz 2001:23).

The relationship between these substances and music has not been studied extensively, but it is known that opioid and monamine secretion is modified in response to music. Goldstein (1980) undertook one of the first studies of the effect of music on opioid secretion. He found that administering the opioid antagonist naxalone affected the frequency, duration and intensity with which "chills" have been experienced in response to music (see also Panksepp 1995; Panksepp & Bernatzsky 2002). The administration of naxalone reduced the chills because it interferes with dopamine production (Levitin 2006:185). It has been thus concluded that the chill response to music is partly controlled by endogenous opioids. Another link to opioids, sound and music is drawn by Panksepp & Bernatzy (2002:137) who mention that the inferior collicus, "an obligatory brainstem way-station for auditory processing", mediates affective processes and has many opioid receptors.

`

The release of opioids is probably responsible for the attachments that are developed to sounds that elicit positive affect.

Dopamine is the main neuromodulator researched in musical experiments. The field of dopamine studies is vast and is rapidly expanding (Knyazev 2007:380). Dopamine, a monoamine modulator, is released by the mesolimbic system of the brain that is involved in arousal, pleasure, and transmission of opioids. Mesolimbic dopamine neuron cell bodies reside in the ventral tegmental area. These cell bodies project to the nucleus accumbens (Kelley & Berridge 2002) where dopamine plays a role in the transmission of opioids. It is the involvement of dopamine in the nucleus accumbens and ventral tegmental area that is the key factor in the rewarding effect of food, sex, drugs of abuse and also music (Menon & Levitin 2005:182). The important role that the nucleus accumbens and dopamine plays in music was unknown until 2005 (Levitin 2006:121). The conventional view that dopamine simply mediates reward is now considered a gross simplification. Berridge & Robertson 1998 (in Knyazev 2007) propose that dopamine is linked to 'wanting' and not 'liking', but this is still debated (Levitin & Menon 2003).

Dopamine furthermore plays a very important role in motor programmes (Herlenius & Lagercrantz 2001). Dopamenergic systems are involved in voluntary motor timing and learning. Motor learning is crucial in producing sensorimotoric-synchronised movements. Panksepp (1998a) suggests that the ancestral basic 'seeking' system is centered on the mesolimbic and mesocortical dopamine circuits. This circuit is involved in the estimation of time and in creating anticipatory states related to rhythmic movements. This may lead to musical expectancies and may be at the heart of the "emotional" impact of music (Panksepp & Bernatzsky 2002:136). The cerebellum and basal ganglia are are also rich in opioid and dopamine receptors and therefore it is highly likely that musicking involving activities such as drumming and dancing would elicit feelings of reward and satisfaction.

Dopamine probably influenced the evolution of long distance running which played a role in rhythmic limb control, as mentioned in Chapter 3. Previc (1999) links dopamine production to thermoregulation during endurance activities such as long distance running and ultra-marathons. Development of bipedalism around 1.6million years ago made long distance running and chase-hunting possible. Chase-hunting involves "continually pursuing an animal in the hot sun until it enters a hyperthermic condition and dies (''chase-myopathy'')" (Previc 1999:319). The physical endurance necessary for chase-hunting requires sweating and rehydration, but crucially, also thermoregulation. Previc (1999: 320) suggests that the need to lower body temperature during

chase-hunting was met by an expansion of dopaminergic systems. Increased dopaminergic innervation would have supplied hominins with superior endurance capability. This change in physical activity and diet (more meat due to more effective hunting) would have influenced further augmented tyrosine production and its conversion to dopamine in the central nervous system (CNS).

The increased dopamine availability also would have affected cognition. Dopamine is involved in working memory, motor planning, cognitive flexibility, abstract reasoning, temporal analysis/sequencing and generativity (Previc 1999). Dopamine bias cells to mainly respond to strong input and therefore focus attention - Le Doux (2002:189) mentions studies that show that the depletion of dopamine in the prefrontal cortex functions as effectively as a complete removal of prefrontal cortex in monkeys. Infusion of dopamine into the prefrontal cortex of young monkeys enhances the working memory capacity and reverses the age-related decrease in working memory.

Two neuropeptides, arginine vasopressin (AVP) and oxytocin (OXT) interact with dopaminergic neurotransmition in reward pathways. Granot *et al* (2004) have found a relationship between musical memory, AVP and OXT and Bachner *et al* (2005) found that these neuropeptides respond to creative dance. Granot *et al* (2004) suggests that arginine vasopressin 1a receptor gene is associated with musical memory in various species. AVP and OXT are involved in social behaviours such as social bonding, regulation of maternal behavior, male courtship, territorial defense and paternal care. It is also concerned with processing and memory for social cues, including facial features, scent, and voice as well as with sensorimotor processes underlying acoustic communication in a number of species. In humans higher levels of OXT are correlated with greater openness to social interactions, less aggression, and more trust. There is also a relationship between AVP and memory. Prior to the analyses by Bachner *et al* and Granot *et al* the effects of AVP have not been studied (Granot *et al* 2004: 313). OXT has been mentioned in musically related researched before. Freeman (1995, 2000a see also Huron 2001) has suggested that OXT underlies the music's capacity to promote group bonding. Granot *et al* (2004:313) conclude "Given the prominent role of vasopressin in social behavior, the preliminary association found in our study between musical memory and vasopressin could serve to support evolutionary accounts postulating a social adaptive role in music, such as mother–infant communication, sexual selection, group cohesion, and even early protolanguage."

`

The biochemical effects discussed in this Section indicate that music making promotes positive states and motivation and this could have been crucial in directing developmental processes and evolution as well. Habitual bipedalism, changed vestibular configuration and respiratory patterns made it possible to run long distances during chase-hunting. This evolved system also made it possible to move rhythmically. Both these activities lead to increased dopamine availability and feelings of reward. It is likely that long distance running, rhythmical movement and dopamine production were linked in the evolutionary origins of music. However, comprehensive understanding of the relationships between aspects of musical activity and biochemical effects is not yet possible, as research into the underlying neurochemistry of musical processing has barely begun and it is largely unknown (Panksepp & Bernatzky 2002:138; Granot *et al* 2004:314). Yet, it is on the level of biochemical studies that advances could be made in understanding how the primary neural pathways (those that are more dependent on genetic mechanisms) that are instantiated in the older medial and ventral pathways, and which utilise neuropeptides (Panksepp 1998) are related to musical experience and action.

6.6 RHYTHMICAL MOVEMENT AND NEURONAL ENTRAINMENT

The capability to move rhythmically rests on the adaptations related to habitual bipedalism and the vestibular system that evolved around 1.6million years ago, as discussed in Chapters 2 and 3. Mithen (2005) and Morley (2003) propose that rhythmic body movement was part of musilanguage, but they do not further analyse its significance. However, Panksepp and Bernatzky (2002:136) observe that rhythmic body movements influence brain rhythm patterns. Benzon (2005: 75) further notes that "…the physical actions – whether dance or vocalization – of different individuals are synchronized on the *same time scale* as that of neural impulses, that of milliseconds." In the following Sections the relationship between rhythmical movement and neural rhythm is explored.

6.6.1 Coupling of bodily and neural rhythms

Large (2000, see also Large & Palmer 2002 in Janata & Grafton 2003:683) discusses the link between rhythmic music making and the neural firing patterns. It is suggested that rhythmic aspects of musicking entrain neural oscillators and that this facilitates synchronization of perception and action with the underlying beat in music (Large 2000, Large & Palmer 2002 in Janata & Grafton 2003: 683). The action of making music couples neural circuits and sensorimotor systems across multiple levels. An individual's rhythmic movement may couple brain rhythms to his movement, as

`

experiments with robots have shown (Eck *et al* 2000). Simple coupling may be synchronous foot tapping to a beat, while more complex coupling would involve dancing or playing a musical instrument (Janata & Grafton 2003).

The basis for the argument that bodily and neural oscillations can be entrained with each other is the phenomenon of coupled oscillators. A pendulum is an example of an oscillator. For example, pairs of pendulum clocks mounted to the same wall would over time synchronize because they influence each other through vibrations in the wall. The entrainment in living systems is more intricate than entrainment of clocks. The concept of coupled oscillation has been applied to the phenomenon of synchronized blinking by fireflies (Strogatz and Steward 1993 in Benzon 2005). In living systems like fireflies there is energy transduction on input (detecting other blinks) and output (generating blinks) and some amplification in between. Verebrate neural systems and human neural systems contain oscillating circuits that exhibit rhythmic behaviour (Freeman 1995; Benzon 2001:53). Benzon (2001, 2005) suggests that attending to one another through musicking results in the nervous system of the performers attuning to each other. Thus, for the duration of musicking nervous systems are physically coupled to one another (Benzon 2001: 47-68). Benzon further notes that entrained musicking activates widespread neural networks and coordinates the temporal flow of neural activity. He (2001:57) builds on a hypothesis of Wallin (1991) who proposed that listening to music influence or entrain one's neural dynamics to the music's dynamics. Benzon emphasises that it is participating in music rather than listening to music that will result in this entrainment.

This proposition holds much promise for understanding the implications of the evolution of musical rhythmical capabilities because it relates bodily movement to neural activity and cognition. Much experimental work still needs to be done in this regard, but the principle that rhythmical sound and action encourage synchronised neural activity is sound. Neurons that fire together in synchrony in widespread brain regions are temporarily bound together, and this coherence of firing, when combined in just the right way across the brain, facilitates various aspects of emotion, motivation and cognition as Knyazev (2007:378) explains: " Growing evidence suggests that different levels of cerebral integration mediated by spatial and temporal synchrony over multiple frequency bands could play a key role in the emergence of percepts, memories, emotions, thoughts, and actions…".

Spatial and temporal synchrony take place over multiple frequency bands. EEG is used to measure the frequency bands of neural activity. The range of frequencies can be divided into global (delta, theta and alpha) and local (beta and gamma) oscillations or modes. The global mode spans

`

relatively large cortical regions and integrate diverse cortical sites. The local modes are higher in frequency than the global modes and lower in amplitude and cover a more limited cortical area. The global and local modes cooperate closely (Knyazev 2007:378). The global modes are discussed here, as they co-regulate dopaminergic systems involved in the motivational, rewarding and cognitive aspects of music.

Delta oscillations at <1 Hz are the slowest and generated in sleep. Delta oscillations are generated cortically and sub-cortically. Among the sub-cortical regions are those associated with the dopaminergic reward system, like the nucleus accumbens (Nacc) and the ventral tegmental area (VTA). These areas create phasic firing that is involved in the release of dopamine (the medial prefrontal cortex and nucleus reticularis thalami are also involved in the delta firing). The phasic dopamine release has been related to feelings of 'intense interest', engaged curiosity' and 'eager anticipation' (Panksepp 1998a). Increased delta activity and dopamine depletion are associated with reward deficiency and craving which have been linked to impulsive, antisocial and aggressive behaviour to achieve relevant rewards (Knyazev 2007:382).

Theta oscillations are generated by a large number of brain structures, including the hippocampus, the 'limbic system' and the frontal regions of the neocortex. Theta activity has been related to learning of movements, but mostly to emotional discrimination in humans. Theta oscillations, along with delta oscillations generate salience detection. Knyazev (2007:383) mentions that the playing of 'alpha' music (Kabuto *et al* 1993) and "ecologically valid sounds" (Shimai 1992) influence brain waves in the theta range. A preponderance of slow wave activity (delta and theta) is linked to impulsivity and is more prominent in for example individuals with Attention Deficit/ Hyperactivity Disorder (ADHD).

The alpha brain rhythms (9-12 Hz) are generated by the anterior and posterior brain regions. These rhythms are linked with inhibitory motor and cognitive processes. Inhibition of impulses is necessary for focussed attention. "Surround inhibition" refers to the phenomenon in which active cortical areas, as evidenced by alpha desynchronization, are surrounded by a "doughnut" of alpha synchronization. This is compatible with Crick's (1984 in Knyazev 2007) spotlight or searchlight of attention hypothesis. Contrary to the 'idling' hypothesis (Knyazev 2007) alpha oscillations are very important in cognitive computation because inhibitory mechanisms are necessary for attention and learning. This rule does not always apply as alpha oscillations may be involved with inhibition of some processes and simultaneously activate other functions. The prefrontal cortex integrates

`

motivational, emotional, sensory, motor and mnemonic information and is essential in suppressing inputs that can interfere with the intended task. Knyazev (2007) suggests that the prefrontal cortex may mediate the reciprocal relationship between alpha, beta and theta oscillations and simultaneously activate alpha-related cognitive operations through inhibition of internal impulses associated with delta and theta activity.

It could not be determined which types of rhythms influence which types of neural oscillations – the only musical examples discussed are those mentioned in Knyazev (2007) in relation to theta waves. Further experiments that assess how musical rhythm and movement influence these global modes would aid in understanding the relationship between music and cognition in evolution. This discussion attempted, in a limited way, to understand how musical rhythmical behavior may influence neuronal firing patterns, but it is also to important to investigate why such entrainment may have been important in the protomusical behaviour. This aspect is discussed below.

6.6.2 The purpose of entrainment

Entrainment through musicking happens at will and is voluntary (Merker 2000: 319-319). Individuals compare their perceptions of their own musical output with their perceptions of the musical output of others. By agreeing to synchronize with each other individuals give up considerable freedom of activity (Benzon 2005: 75). Freeman (2000 a,b) has put forward the argument that internal representations of the world need to be confirmed by other conspecifics through entrained actions because all knowledge is constructed and relative. Freeman is one of the pioneers in interpreting the nervous system's functioning in terms of chaos, complex systems and nonlinear dynamics (Benzon 2001:56), but his ideas are not always mainstream (Raushecker 2002:1032). According to Freeman, knowledge is not 'out there' to be objectively perceived and processed - it is dynamically created by the brain. He (Freeman 1995, 1999, 2000 a,b) experimentally observed electrical activity patterns of auditory and somatosensory pathways during experiments involving conditioned stimuli. Such neuronal activity is characterized by aperiodical wave patterns that last for only about a tenth of a second. The content of the stimulus is related to the spatial pattern of amplitude modulation (AM) of the common chaotic waveforms that act as carrier (Freeman 2000:413). These AM patterns converge and are combined in the limbic system and are integrated in gestalt patterns. The new gestalt pattern replaces the original neuronal stimulus pattern. There is no relationship between the original neural firing and the constructed neural firing. Freeman interprets this as an indication that the meanings of stimuli are constructed in terms of past

`

experience. There is thus no way in which two individuals' experience of the external world can be the same. Knowledge can be made public and validated by means of intentional action between individuals. Shared actions are therefore necessary to create and establish mutual understanding and trust.

Mithen (2005: 216-217) discusses and accepts Freeman's ideas on "trust" and its implications for group bonding. Trust is defined as "…an implicit expectation and faith in the predictability of the behaviour of those to whom one has committed oneself by a transformation of the self" (Freeman 2000:420). Freeman's notion of 'trust' is similar to the intersubjectivity (Trevarthen & Aitken 2001) and mutuality (Miall & Dissanayake 2003). The strongest basis for cooperation is rhythmically repeated motions, because they are predictable, can be anticipated and moved to. Through such activity individuals' responses can be tightly entrained (Benzon 2001, 2005). When nervous systems are coupled through interactional synchrony, a uniquely human social space is created (Benzon 2001:28). Furthermore, there is a feeling of reward and satisfaction as the brain, through activity in the cerebellum synchronises neural oscillators to the pulse of the music (Levitin 2006:187). This link between neurophysiological mechanisms and social processes through which musicking facilitates the coupling of human behaviour in time through sound and action is convincing (Cross 2001b). It is interesting that Williams (1967 in Freeman 2000:420) extends cooperation through rhythmically repeated motions to apes, even though they lack the sense of rhythm of humans.

Benzon (2001:175) proposes that the origin of rhythmical movement in hominins may have been in imitation. Imitation of animal calls is for Benzon (2001:171, 173) one of the adaptations that allowed *Homo ergaster,* at 1.6million years ago to survive in the steppes, or savanna environment. For him, imitation of movement and behaviour follow logically from here. Benzon suggests that the imitation of animals' rhythms and making of stone tools may have be what instigated bodily rhythmic behaviour. He proposes that vocal mimicry was followed by "a Gestalt switch or self-organized change in dynamics, in which rhythmically coordinated group musicking emerges from vocal mimicry and associated activities" (:172). He sketches a scenario in which a band of protohumans become 'edgy' and then, "….somehow, the group members begin stomping their feet in the same rhythm and start to vocalize wildly, each singing his or her own line yet all somehow managing to blend together in a fine raucous mix. They do this for an hour or so, and the anxiety gradually dissipates as the rhythmic actions recruit more and more neural circuits into the flow,

`

dissolving the neural traffic jams" (Benzon 2001:88). There are a number of problems with this scenario: It is unlikely that the source of rhythmic movement was rhythmic stone tool production. Even though stone tool making requires a certain degree of rhythmicity, it is far removed from entrained rhythmicity (see also Cross 2001b). The most problematical part of Benzon scenario is that group members 'somehow' began joint rhythmic behaviour. Previc's research on long-distance running, the regulation of temperature and dopamine secretion is a more credible source of rhythmic musical movement. These movements may have been initially associated with grunt vocalisations automatically expressed with effortful activity (compare McCune *et al* 1996).

Rhythmic movement accompanied by rudimentary melody may have provided the social cement to bond a group, but it could also have served as mnemonic device to build the first symbolic networks (Benzon 2005). This suggestion of Benzon is not unlikely. Intentional rhythmical entrainment may have been crucial in achieving synchronised coherent brain rhythms necessary to think in symbols. Symbolic conceptualization makes use of various areas of the brain simultaneously - the effective communication between these areas determines the ability to sustain a coherent message. To be able to plan thoughts, or patterns of neural impulses (Calvin & Bickerton 2000:201), must be held in the brain long enough to allow assembling and reassembling of the neural impulses. Creating a sentence requires transmitting signals over long distances within the brain without losing the coherent functioning. The ability of cell assemblies or groups of cells to fire simultaneously or in resonance, strengthens the signal and prevents signal distortion. The mechanisms that evolved through musicking could have been instrumental in creating the neural synchronisation and coherence necessary to think in symbols.

6.7 DISCUSSION

It is extraordinarily complex to unravel music's psychological effect as there is no constant relationship between a certain aspect of music and a certain affective state as measured by self-report. There is no point in playing music to a passive listener who assesses his own emotional assessment using standard psychological tools when the experimental aim is to uncover universal human reactions to music. Experimental situations and tasks should be realistic - they should have 'ecological validity' (Clarke 2005). The strength and speed of the response that music elicits (Huron 2001:44) is not an indication of a meaningful intrinsic relationship between a musical element and an emotion. Perhaps the strong and immediate reactions to music point to the importance of music in social life. It is not the particular emotional reaction that is adaptively significant; it is the

`

centrality of music as an emotionally meaningful social practise that is important. Humans utilise music to achieve positive emotional states, much like infants and infants utilise music-like communication to create interpersonal relationships. Humans use music as "communicative offering to influence our feelings in a re-creative process" (Grewe *et al* 2007:313).

This chapter investigated whether there are biological "emotional" reactions to musical stimuli that are universally present and therefore significant in the study of the evolutionary origins of music. Bartlett's (1996:438) observation, that the possibility of finding a direct relationship between a physiological response and the psychological manifestation of that response is remote (Bartlett 1996:348), has been confirmed. The utilisation of 'music' as stimulus in many of the experiments discussed here is problematical. Music for example is classified as stimulative, neutral or sedative, but there are no general rules for or consensus on which musical pieces belong to certain categories. Also, researchers rarely use the same piece of music. Different variables such as rhythm, pitch and loudness are used to measure the same physiological response, and sample sizes vary greatly (Bartlett 1996:349). It is especially problematic to determine whether the measured physiological response are actually related to the musical stimulus, or due to specific or general affective changes evoked by the music.

The study of biochemical reaction to music holds much promise for understanding its evolutionary origins, but at this time there seems to be more questions than answers. Nevertheless, increased dopamine excretion, related to habitual bipedalism, long distance running, chase-hunting and rhythmical movement that evolved around 1.6million years ago was most probably a pivotal evolutionary significant occurrence. Rhythmical movement further affects brain rhythms – bodily and neuronal oscillations are synchronized with each other through entraimed rhythmical movement. Through entrained rhythmical movement hominins could have engendered "trust", "mutuality" or "intersubjectivity" and in this way promoted group cohesion. Protomusic consisting of entrained rhythmical whole body movements associated with grunt-like vocalisations, would therefore have been adaptive. The discussion in this chapter provides support for the hypothesis, mentioned Chapter 3, Section 3.4.2, that protomusic functioned to strenghten social bonds with conspecifics to promote group-cohesion.

`

CHAPTER 7: CONCLUSIONS

The purpose of this study was to investigate the evolutionary origins of music. The available literature on the origins of music almost entirely addresses the subject from the perspective of musilanguage. Therefore two recent multidisciplinary publications, a book by Mithen (2005), "The Singing Neanderthal", and a doctoral dissertation, "The evolutionary origins and archaeology of music", by Morley (2003) were used as starting point to identify suitable criteria for further investigation into the origins of music. Whether music and language have common origins, as suggested by the musilanguage theory, has not been investigated, as this would have required a comprehensive study that demands expertise from for example, a linguistic point of view. A biomusicological approach was taken in this study of the origins of music. To define music perspectives from musicology, comparative biology, developmental psychology, anthropology and the science of movement were incorporated. The following elements of music have been identified:

a) Music is an ententional action that encompasses the whole body. Most known peoples' musics are integrated with dance and action, often in ritual context.

b) Music relies in part on the capability to produce complex and learned vocalizations. The possibility that complex learned sound could have been produced instrumentally cannot be excluded, but this aspect was not investigaged in this biomusical study. Musically produced sound has a specific configuration of pitch and timbre.

c) Music involves the ability to predict where a rhythmical beat will fall, and on this basis actions can be entrained to the beat. This makes dancing possible. Dancing involves rhythmical action of the whole body and is integrated with the production of musical sound.

Music was thus defined as "an intentional action in which complex, learned vocalizations (and/or instrumentally produced sound) are combined with the movement of the body in synchrony to a beat. Whole body movements are involved."

Four research questions were posed in the introduction. Sections 7.1 – 7.5 discuss the responses to these questions while in Section 7.6 the status of the research hypothesis is discussed.

7.1 THE FOSSIL EVIDENCE FOR THE PRODUCTION AND PERCEPTION OF MUSIC

The first research question posed was: "Which biological aspects leave traces in the fossil record that can be reliably related to music production and perception, and what implications do these finds have for the evolutionary origins of music? This question was addressed in Chapters 2 and 3 by describing the biological adaptations and fossil correlates of the capabilities involved in vocal production, hearing, respiration and rhythmical movement or dancing. The adaptations that make it possible to produce vocal musical sounds, the descent of the larynx and changes to the respiratory system, and their related fossil traits, as presented by Mithen and Morley, were investigated. The descent of the larynx in humans has lead to an elongated vocal tract. This makes it possible for humans to produce a wider range of sounds than other primates. The timing of the descent of the larynx has been tied to the fossil evidence for bipedalism (Mithen & Morley), the flexion of the basicranium and morphology of the hyoid bone (Morley) and the size of the hypoglossal canal (Mithen). It has been shown in Chapter 3 that the initial descent of the larynx predates the split between chimpanzees and the Homo lineage 6 million years ago and it is unknown when a second phase of descent took place. Therefore the fossil evidence for bipedalism is not relevant to the descent of the larynx. As the flexion of the basicranium, morphology of the hyoid bone and the hypoglossal canal cannot be reliably related to the position of the larynx, there are no reliable fossil indicators for the timing of the descent of the larynx. The fossil record thus cannot inform on the timing of the ability to perform a larger-than-primate range of vocalizations. None of the fossil evidence discussed by Mithen and Morley could be related to vocal production with confidence.

In contrast, the fossil evidence for the adaptation related to the evolution of the refined neural control of the breath, unique to humans, is an important indicator of the development of vocal music. In Section 2.4.2 it was explained that human singing requires the use of all the major respiratory muscles. The size of the thoracic vertebral column is an indication of the number of nerve cells that control the breath via the intercostal and the abdominal muscles (MacLarnon & Hewitt 2004). A bigger thoracic canal indicates that more thoracic spinal nerves were available to control respiration. The thoracic vertebral canal expanded some time after 1.6 million years ago and attained modern size by 600 000 years ago. Mithen and Morley related the size of the thoracic canal to the ability to speak, but not to singing. However, Fitch (2006a:196) emphasises that the although the studies of MacLarnon & Hewitt have been undertaken in the context of speech "… the changes they have documented are equally, if not more, relevant to song…These data are thus consistent

`

with the hypothesis that song evolved either before, or simultaneously, with speech." The hypothesis can therefore be formulated that by 600 000 years ago, archaic humans evolved the respiratory control necessary to sing.

The evolution of habitual bipedalism at 1.6 million years ago, together with anatomical adaptations for long distance running (as discussed in Chapter 6), is related to a number of developments. One of these is the adaptation of the vestibular system of the inner ear. The research of Spoor and colleagues has shown that the different way in which hominins moved around 1.6 million years ago affected the orientation of the three semi-circular canals of the vestibular system. The vestibular system is involved in the proprioceptive awareness of the body and, importantly, perhaps rhythmic movement. There is a need for the evolutionary relationship between the vestibular system and rhythmic movement to be developed and clarified. There is a promising proposition (Phillips-Silver & Trainor 2008:94) that "... the tie between music and movement is concrete …and that, indeed, music may have evolved from physical movement (e.g. Todd *et al*, 2007; Trainor 2007)". A number of studies have been undertaken (e.g. Phillip-Silver & Trainor 2005, 2007, 2008) that show a clear interaction between movement and auditory rhythm perception in infants and adults. Mithen and Morley made reference to the vestibular system and its relationship to bipedalism, but did not tie it specifically to the evolution of the capability to move rhythmically. There is scope to develop an extensive hypothesis on how movement and rhythm could have interacted in the evolutionary origins of music. The evolution of the adaptations relating to bipedalism, long distance running and the vestibular system were important in the development rhythmically entrained movement of the whole body, or dancing. It is plausible that, by 1.6 million years ago, hominins were able to dance.

The lowered larynx would have made an extended range of sounds possible, but it is necessary to explain how these types of vocalization could have developed into human singing that involves conscious manipulations and novel recreations of vocalizations. The combination of laryngeal grunts with rhythmical movements may have preceded the development of the complex singing in protohumans. The grunt is an ideal candidate for a transitionary vocalization that could have been used in association with rhythmic movements. Grunts, as discussed in Chapters 2 and 4, occur consistently in many species in association with effort and movement. Grunts eventually develop into intentional vocalisations that are ritualized as communicative signals in various species (McCune *et al* 1996). It may have been the act of combining vocalization with the newly evolved ability to move rhythmically that allowed flexible vocalizations to develop. It is conceivable that

`

vocalizations similar to extant ape-calls were part of hominins' vocal repertoire. Vocal learning associated with grunt vocalizations could have been the first simple vocalizations combined with rhythmical movement. The ability to learn vocally could have been transferred to more extensive ape-calls to produce the first complex learned song. There is limited fossil evidence for the evolution of the capabilities underlying music. The ability to move rhythmically may have evolved around 1.6 million years ago, before the capability to control respiration which was in place by 600 000 years ago. Protomusic probably involved rhythmical movement coupled to grunt-like vocalizations.

7.2 BRAIN EVOLUTION AND THE ORIGINS OF MUSICALITY

Mithen and Morley propose that the evolution of modularisation and lateralization underlie modern music capabilities. Chapter 4 assessed their evolutionary paths for the brain and music to determine, as articulated in the second research question, whether there is any evidence from the brain sciences that can inform on the evolutionary origins of music. The review of neuroscientific studies on brain evolution established that the facts relevant to the origins of music are few. The evolutionary design of the brain is such that prespecified genetic information for connectivity is minimal. The brains of all organisms are bio-social organs in which the input from the environment is integral. The wiring of the brain is a Darwinian process and new functions arise epigenetically. This view of brain evolution is contrary to the evolutionary paths proposed by Mithen and Morley. Modularisation and lateralisation are not specified genetically. It is relevant that human brains are unique in their relative brain to body size, and in the relative proportions of the pre-frontal cortex and the cerebellum.

At 600 000 years ago there was an increase in brain size relative to body size. This enlargement of brain size independent of body size must have brought about changes in brain organization. Larger brains with more neurons require cell assemblies to be organised differently to effectively maintain communication between the different parts of the brain, as discussed in Chapter 4. From the information presented in Chapter 6, it can also be inferred that the brain enlargement must have had implications for the types of synchrony necessary for coherent functioning and the secretion of neurotransmitters and modulators. The larger prefrontal cortex and cerebellum necessitated different connectivity patterns. There is no simple relation between a brain area and a cognitive function, but in the main the prefrontal cortex regulates the executive functions such as working memory and symbolic language and perhaps modern music, while the cerebellum manages

`

movement patterns and relates motor patterns to cognition. This is a further indication that the argument for the modular view of brain and musical origins cannot be supported.

In higher organisms like primates with complex behaviour patterns, behaviour interacts with the body or anatomical basis of an organism to steer evolution in a certain direction. This is known as Baldwinian evolution. Even though the process is contested, it remains highly probable that behaviours drove the direction of evolution in *Homo* (Deacon 2003). It is probable that protomusic drove the direction of evolution towards the development of music as an intentional act in which complex learned vocalization and rhythmical movement were combined.

7.3 INFANT-DIRECTED COMMUNICATION AND INNATE MUSICAL CAPABILITIES

Infant-directed communication involves singsong vocalizations of caregivers combined with rhythmically and vocally synchronous actions between them and infants. The musicality of infant-directed communication has led to the hypothesis, as discussed by Mithen and Morley that infants are born with inherent musical capabilities. Therefore the third research question was directed at establishing whether infants are born with innate musical capabilities and whether Infant-directed communication is relevant to the evolutionary origins of music. The investigation of this hypothesis necessitated determining whether the development of capabilities that infants are born with starts *in utero*. It was found that the foetus is motivated to actively exploit the benefits of the environment *in utero*. This determines the motivational and attentional predispositions of the foetus and newborn. From the earliest stages the foetus learns through bodily actions and therefore the newborn is preconditioned to use his bodily actions and perceptive organs to communicate intent.

The literature review in Chapter 5 has shown that, as Mithen and Morley suggest, the infant is born with some rudimentary musical skills – he has superior hearing abilities and sensitivity for melody, sensitivity for spotting changes in temporal sequences and an ability to imitate vocally and bodily. Infants are not born with the capability to entrain their movements to an external beat, but they are extraordinarily sensitive to changes in tempo and they can adjust their sucking rate to tempos that are similar to or faster than their endogenous tempo. These capabilities are not musical in the sense of an intentional act, but they predispose human infants to learn in a musical way and to prefer musical communication.

The human infant is unique amongst primates in terms of the degree to which he is specifically receptive to the subjective states in other people. This intersubjectivity is described as a specialized,

innate, 'human-environment-expectant' and socially regulatory capability (Trevarthen & Aitken 2001:4). Infants are specifically motivated, beyond instinctive behaviours, to communicate their need for social interaction. The infants' innate capability to imitate a broad range of whole-body actions, facial expressions and vocalizations are crucial in their development. Infants learn through rhythmical movement and imitation through a process of autopoiesis. The reciprocal relationship between for example movement and rhythmical learning, between vestibular stimulation and gross motor abilities, and between rhythmicity, imitative vocalisation and the acquisition of language demonstrate the usefulness of the evolved musical system.

Mithen's and Morley's ideas on the innate musical capabilities in infants have been confirmed, but the aspects emphasised in this study are somewhat different. The motivational state and imitative abilities are given more prominence but the conclusion, in principle, is the same. The infant is born with inherent rudimentary musical capabilities. Infant-directed communication is thus relevant to the study of the origins of music. The tendency to mimic extensively and to synchronize vocalizations and movements with others to achieve emotional convergence or intersubjectivity may reflect the essence of music in the evolutionary past.

7.4 THE BIOPSYCHOLOGICAL EFFECT OF MUSIC

It is clear that music has a strong "emotional" impact, but researching a nebulous concept such as "emotion" is problematical. "Emotion" is often studied from a psychological perspective, but such reactions to music are idiosyncratic, individualistic and determined by cultural preferences. The psychological investigation of musical emotions has little relevance to the study of the origins of music. For this reason the emotional effect of music has been investigated from a biological point of view in order to answer research question 4: "Are there "emotional" responses to music that are evolutionary meaningful?"

Mithen and Morley both see the emotional effect of music as evolutionarily important, but do not elaborate in depth on why this could be so. To further investigate music's potent "emotional" effect, the influence of music on the immune and neuroendocrine systems was investigated. It was found that music making promotes positive states and motivation and that this can be related to certain biochemical reactions that may have evolutionary significance. The study of the biochemical effect of music and musicking is in its infancy, but rhythmical movement and musicking have been associated with increased immunological competence and the release of the neuromodulator

dopamine. Rhythmical musical activities engage the cerebellum and the sub-cortical parts of the brain where these substances are released and this has an effect on mood regulation, feelings of anticipation, reward and satisfaction. It is in this sense that music has an emotional impact. More research needs to be undertaken to unravel the correlation between aspects of musical activity and biochemical effects.

Another biological reaction to music that has been investigated in depth here was the effect of rhythmical movement on the neuronal system. The rhythmic aspects of musicking entrain neural oscillators and this facilitates syncronization of perception and bodily action with the rhythm. This entrainment result in the coupling of individuals' nervous systems and in the process a uniquely human social space is created. Knowledge of the world is constructed and 'deeply personal' and individual. Therefore there is a need for individuals to bridge the "solipsistic gulf" (Freeman 2000:411). Trust, as Freeman and Benzon explain, can be achieved through musical rhythmic behavioural activities that lead to the dissolvement of the solipsistic barrier by the secretion of neurohormones such as oxytocin. This may have been the mechanism involve in fulfilling the need for increased social bonding. This analysis supports the hypothesis put forward by McNeill (1995), Benzon (2001), Freeman (2000), Morley (2003) and Mithen (2005) that the purpose of musical activity was to create group cohesion and trust. This could have been achieved by protomusic that may have consisted of rhythmically entrained movement combined with simple learnt grunt-like vocalisations that were gradually replaced by more complex learned song.

7.5 EVALUATING THE RESEARCH HYPOTHESIS

The biological basis of music was investigated in an effort to understand its origins and rudimentary form. The musilanguage hypotheses as put forward by Mithen and Morley provided an exceptional basis for such an investigation. Four research questions were formulated on the basis of their work and the following hypothesis was set:

The propositions made by Mithen and Morley about the evolutionary origins of music are open to criticism and they can be revised in the light of relevant literature. The propositions in question are that there is sufficient evidence from the fossil record to reliably infer the evolutionary origins of the production and perception of music; that the neuro-scientific literature supports the notion that the evolution of lateralisation and modularisation of the brain underlies musical capabilities; that infants have innate musical capabilities; that there are certain "emotional" responses to music that are meaningful from an evolutionary perspective. These propositions indicate that the capabilities

`

for human music making, in the form of entrained movement and complex learned vocalisation, evolved between 1.6 million years ago and 50 000 years ago.

This hypothesis can be accepted for the following reasons:

There is limited evidence from the fossil record to infer the evolutionary origins of the production and perception of music. The firmest evidence for the evolution of musical abilities relates to the evolution of rhythmically entrained movement and this provides the most promising avenue for future research on the origins of music. The review of the evolutionary biomusicological and palaeoanthropological literature has shown that there is more uncertainty around the evolution of vocal capabilities than evident from Mithen's and Morley's accounts.

The neuro-scientific literature does not support the notion that the evolution of lateralisation and modularisation of the brain underlies musical capabilities. Pre-specified genetic information for connectivity and brain organisation is negligible and the wiring of the brain is largely an epigenetic process. The relative size of the pre-frontal cortex and cerebellum is most probably genetically determined and involved in the evolution of musical capabilities.

The infant is born with rudimentary musical capabilities in terms of superior hearing and sensitivity for melody and changes in temporal sequences. These skills together with the human infant's unique vocal and bodily imitative capabilities and receptiveness for the subjective states of others predispose the infant to learn in a musical way and to prefer musical communication. It is in this sense that infants are innately musical.

Psychological "emotional" response to music has little relevance to its evolutionary origins, but there are biological "emotional' reactions that are relevant. The biological "emotional" response to music in terms of biochemistry and neuronal synchrony is most probably significant for the evolutionary origins of music, but research in this field is still in its infancy.

It is plausible that protomusic consisting of rhythmically entrained body movements associated with simple vocalizations evolved in the window of time between 1.6 million and 600 000 years ago, and not between 1.6 million and 50 000 years ago as suggested in Mithen's work. The adaptations associated with bipedalism, long distance running and the vestibular system that evolved around 1.6 million years ago could have been utilized for entrained rhythmical musicking. It is plausible that the first vocalisations that were combined with rhythmic movements were grunt-like vocalizations.

`

The evidence discussed in this thesis is not strong enough to ascertain when complex learned vocalization developed and to therefore infer the origins of modern music. Modern music would have evolved when intentional synchronisation of body movements was combined with consciously created learned and complex vocalizations. The fossil evidence shows that the adaptations underlying modern breathing patterns were in place by 600 000 years ago. In theory this would have made the extensive breathing control that allows learned and complex vocalization, for example singing, possible.

Protomusic, developing after 1.6 million years ago, may thus have consisted of entrained rhythmical whole body movements initially combined with grunts. Rhythmic entrainment of body movement may have been coupled to these simple vocalisations to achieve synchronous mind states. The capability to control the breath voluntarily was possible by 600 000 years ago, hypothetically enabling complex vocalization. Although the evidence discussed in this thesis is not strong enough to infer the origins of modern music, it is plausible that it had its roots in the Middle Pleistocene.

`

BIBLIOGRAPHY

Abdelhak, S., Kalatzis, V., Heilig, R., Compain, S., Samson, D., Vincent, C., Levi-Acobas, F., Cruaud, C., Le Merrer, M., Mathieu, M., König, R., Vigneron, J., Weissenbach, J., Petit, C. & Weil, D. 1997. Clustering of mutations responsible for branchio-oto-renal (BOR) syndrome in the eyes absent homologous region (eyaHR) of EYA1. *Human Molecular Genetics* 6: 2247–2255.

Abeles, H.F. & Chung, J.W. Responses to Music. In D.A. Hodges (ed) *Handbook of Music Psychology*. San Antonio: IMR Press, 285-342.

Abrams, B. 2001. Music, cancer and immunity. *Clinical Journal of Oncology Nursing* 5(5):222-224.

Ader, R. 2000. On the development of psychoneuroimmunology. *European Journal of Pharmacology* 450:167-176.

Adolphs, R. 2002. Neural systems for recognizing emotion. *Current opinion in neurobiology.* 12:169-177.

Aiello, L.C. 1996a. Terrestriality, bipedalism and the origin of language. In W.G. Runciman, J. Maynard-Smith & R.I.M. Dunbar (eds) *Evolution of social behaviour patterns in primates and man*. Oxford: Oxford University Press, 269-90.

Aiello, L.C. 1996b. Hominine preadaptations for language and cognition. In: P. Mellars. & K. Gibson (eds) *Modelling the early human mind.* Cambridge: McDonald Institute Monographs, 890-99.

Aiello, L.C. & Dunbar, R.I.M. 1993. Neocortex size, group size, and the evolution of language. *Current Anthropology* 34:184-193.

Alcock, K. J., Passingham, R. E., Watkins, A. J., & Vargha-Khadem, F. 2000. Pitch and timing abilities in inherited speech and language impairment. *Brain and Language* 75: 34–46.

Aldridge, D. 1993. Music therapy research 1: a review of the medical research literature within a general context of music therapy research. *The Arts in Psychotherapy* 20:11-35.

`

Al-Qahtani, N.H. 2005. Foetal response to music and voice. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 45:414-417.

Anderson, J.W. 2005. The natural ways of musicality, language and self-regulation in infants. *Psychomusicology* 19:41-59.

Annett, M. 1985. *Left, Right, Hand and Brain: the Right Shift Theory.* London: Lawrence Erlbaum.

Aragon, D., Farris, C. & Byers, J.F. 2002. The effects of harp music in vascular and thoracic surgical patients. *Alternative therapies in health and medicine* 8(5):52-60.

Avorgbedor, D.K. 2008. Review of Mithen, S. 2005. The Singing Neanderthals, The origins of Music, Language, Mind and Body. London: Weidenfeld and Nicolson. *Empirical Musicology Review* 3(1):22-35.

Bachner, R., Christian, D., Zohar, A. H., Constantini, N., Lerer, E., Hoch, S., Sella, S., Nemanov, L., Gritsenko, I., Lichtenberg, P., Granot, R. & Ebstein, P. 2005. AVPR1a and SLC6A4 gene polymorphisms are associated with creative dance performance. *PLOS Sciences* 1(3): 394–403.

Balaban M.T., Anderson, L.M. & Wisniewski, A.B. 1998. Lateral asymmetries in infant melody perception. *Developmental Psychology.* 4:39–48.

Bartlett, D.L. 1996. Physiological Responses to Music and Sound Stimuli. In D.A. Hodges (ed) *Handbook of Music Psychology.* San Antonio: IMR Press, 343-385.

Baruch, C. & Drake, C. 1997. Tempo discrimination in infants. *Infant Behavioural Development.* 20**:** 573–577.

Beck, R. J., Cesario, T. C., Yousefi, A. & Enamoto, H. 1999. Choral singing, performance perception, and immune system changes in salivary immunoglobulinA and cortisol. *Music Perception* 18: 87–106.

Beek, P.J., Peper, E. & Daffertshofer, A. 2000. Timekeepers versus nonlinear oscillators: how the approaches differ. In P. Desain & L. Windson (eds) *Rhythm perception and production* Lisse: Swets & Zeitlinger, 9-33.

`

Benasich, A.A. & Tallal, P. 2002. Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural and Brain Research* 136:31-49.

Bendor, D. & Wang, X. 2006. Cortical representations of pitch in monkeys and humans. *Current Opinion in Neurobiology* 16:391-399.

Benzon, W.L. 2001. *Beethoven's Anvil: Music in Mind and Culture.* Oxford: Oxford University Press.

Benzon, W.L. 2005. Synch, Song and Society. Review of Mithen, S. 2005. The Singing Neanderthals, The origins of Music, Language, Mind and Body. London: Weidenfeld and Nicolson. *Human Nature Review* 5:66-86.

Bergeson, T.R. & Trehub, S.E. 2006. Infant's perception of rhythmic patterns. *Music Perception* 4:345-360.

Berridge, K.C. & Robinson, T.E. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research & Brain Research Review* 28: 309–369.

Besson, B. & Friederici, A. 2005. Part II: Language and Music – a comparison. *Annals of the New York Academy of Sciences* 1060:57-58.

Bickerton, D. 2007. Language evolution: a brief guide for linguists. *Lingua* 117:510-526.

Bispham, J. 2006a. "Music" means nothing if we don't know what it means. Review of Mithen, S. 2005. The Singing Neanderthals: The Origins of Music, Language, Mind and Body. London: Weidenfeld and Nicolson. *Journal of Human Evolution* 50:587-593.

Bispham, J. 2006b. Rhythm in music: What is it? Who has it? And why? *Music Perception* 24(2):125-134.

Bittman, B., Berk, L.S., Felten, D.L., Westengard, J., Simonton, C., Pappas, J. & Ninehouser, M. 2001. Composite effects of group drumming music therapy on modulation of neuroendocrine-immune parameters in normal subjects. *Alternative therapies in Health and Medicine* 7(1): 38-47.

`

Blacking, J. 1973. *How musical is man?* Seattle: University of Washington Press.

Blalock, E.J. 2002. Harnessing a neural-immune circuit to control inflammation and shock. *Journal of Experimental Medicine* 195(6):F25-F28.

Blood, A.J. & Zatorre, R.J. 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences* 98(20):1181-11823.

Blood, A.J., Zatorre, R.J., Bermudez, P & Evans, A.C. 1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience* 2(4): 382-387.

Bobin-Bègue, A., Provasi, J., Marks, A. & Pouthas, V. 2006. Influence of auditory tempo on the endogenous rhythm of non-nutritive sucking. *Revue Européene de Psychologie Appliquée* 56:239-245.

Bocquet-Appel, J.P. & Arsuaga, J.L. 1999. Age distributions of hominid samples at Atapuerca (SH) and Krapina could indicate accumulation by catastrophe. *Journal of Archaeological Science* 26:327-37.

Boone, J.L. & Smith, E.A. 1998. Is it evolution yet? A critique of evolutionary archaeology. *Current Anthropology 39*, supplement: 141-173.

Botha, R. *in press*. On musilanguage/"Hmmmmm" as an evolutionary precursor to language. *Language and Communication*.

Bouissac, P. 2004. How plausible is the *motherese* hypothesis? Commentary on Falk, D. Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences* 27:506-507.

Bramble, D.M. & Lieberman, D. E. 2004. Endurance running and the evolution of Homo. *Nature* 432:345-52.

Bräuer, G., Yokoyama, Y., Falquères, C. & Mbua, E. 1997. Modern human origins backdated. *Nature* 386: 337.

`

Brown, S. 2000. The musilanguage model of music evolution. In N.L. Wallin, B. Merker & S. Brown. (eds) *The origins of music*. Cambridge MS: MIT Press, 271-300.

Brown, S. 2006. Are music and language homologues? *Annals of the New York Academy of Sciences*. Xx:372-374.

Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K. & Jancke, L. 2000. Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Research Cognitive Brain Research* 9: 227–238.

Buller, D.J. & Hardcastle, V.G. 2000. Evolutionary psychology, meet developmental biology: against promiscuous modularity. *Brain and Mind* 1: 307-325.

Burling, R. 2004. Prosody does not equal language. Commentary on Falk, D. 2004. Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences* 27:509.

Byrne, R.W. & Bates, L.W. 2007. Brain Evolution: when is a group not a group? *Current Biology* 17(20):R883-R885.

Cabanac, M. 2002. What is emotion. *Behavioural Processes* 60:69-83.

Calvin, W. H. 1996. *How Brains Think*. London: Weidenfeld and Nicholson.

Calvin, W. & Bickerton, D. 2000. *Lingua ex Machina: reconciling Darwin and Chomsky with the Human Brain*. Cambridge, MA: The MIT Press.

Campbell, M. & Greated, C. 1987. *The Musician's Guide to Acoustics*. London: Dent and Sons.

Carey, J. & Amin, V. 2006. Evolutionary Changes in the Cochlea and Labyrinth: Solving the Problem of Sound Transmission to the Balance Organs of the Inner Ear. *The Anatomical Record part A* 288:482-490.

Carrier D. 1984. The energetic paradox of human running and hominid evolution. *Current Anthropology* 25:483–495.

Carterette E.C. & Kendall R.A. 1999. Comparative music perception and cognition. In: D. Deutch (ed) *The psychology of music*. New York: Academic Press, 725–791.

`

Christensen-Dalsgaard, J. 2004. Music and the origin of speeches. *The Journal of Music and Meaning 2,* section 2.

Clark, A. 1999. An embodied cognitive science? *Trends in Cognitive Sciences* 3(9):345-351.

Clark, D.L., Kreutzberg, J.R. & Chee, F.K.W. 1977. Vestibular stimulation influence on motor development in infants. *Science* 196(4295): 1228-1229.

Clark, A.G., Glanowski, S., Nielsen, R., Thomas, P.D., Kejariwal, A., Todd, M.A.,Tanenbaum, D.M., Civello, D., Lu, F., Murphy, B. Ferriera, S., Wang, G.,Zheng, X., White, T.J., Sninsky, J.J., Adams, M.D. & Cargill, M. 2003. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* 302: 1960–1963.

Clarke, E. 2005. *Ways of listening: An ecological approach to the perception of musical meaning.* Oxford, GB: Oxford University Press.

Clegg, M. & Aiello, L.C. 2000. Paying the price for speech? An analysis of mortality statistics for choking on food. *American Journal of Physical Anthropology Supplement* 30: 126.

Clynes, M. 1978. *Sentics: the Touch of the Emotion.* New York: Doubleday.

Condon, W.S. 1986. Communication: Rhythm and Structure. In J.R. Evans & M. Clynes (eds) *Rhythm in Psychological, Linguistic and Musical Processes*. Springfield, IL: Charles C Thomas, 55-78.

Conroy, G., Weber, G., Seidler, H., Recheis, W., Zur Nedden, D. & Mariam, J.H. 2000. Endocranial capacity of the Bodo cranium determined from three-dimensional computed tomography. *American Journal of Physical Anthropology* 113:111-118.

Cook, N. 2001. Theorizing musical meaning. *Music Theory Spectrum* 23:170-195.

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

Corballis, M.C. 1998. Cerebral asymmetry: motoring on. *Trends in Cognitive Science* 2:152-157.

Cosmides, L. & Tooby, J. 1994. Origins of domain specificity: The evolution of functional organization. In L.A. Hirschfeld & S.A. Gelman (eds) *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press, 85–116.

`

Craig, D.G. 2005. An exploratory study of physiological changes during "chills" induced by music. *Musicae Scientiae, IX,* 273-287.

Crawford, J.R., Garthwaite, P.H. & Gray, C.D. 2003. Wanted: fully operational definitions of dissociations in single-case studies. *Cortex* 39:357-370.

Crick, F. 1984. Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of National Academy of Sciences of the USA* 81: 4586–4590.

Cross, I. 1999. Is music the most important thing we ever did ? Music, development and evolution. In Suk Won Yi (ed) *Music, mind and science*. Seoul: Seoul National University Press, 10-39.

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

Cross, I. 2001b. A meeting of musicking minds. Review of Benzon, W.L. 2001. Beethoven's Anvil: Music in Mind and Culture. Oxford: Oxford University Press. *The Dana Foundation* [http://www.dana.org/news/cerebrum/detail.aspx?id=1766 accessed 6 August 2008]

Cross, I. 2003. Music and Evolution: consequences and causes. *Contemporary Music Review* 22(3): 79-89.

Crow, T.J. 2004. Cerebral symmetry and the lateralization of language: core deficits in schizophrenia as pointers to the gene. *Current Opinion in Psychiatry*. 17(2):97-106.

D'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A., Soressi, M. Bresson, F., Maureille, B., Nowell, A., Lakarra, J., Backwell, L. & Julien, M. 2003. Archaeological evidence for the emergence of language, symbolism, and music – an alternative multidisciplinary perspective. *Journal of World Prehistory* 17(1):1-70.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. 2001. A developmental study of the affective value of tempo and mode in music. *Cognition* 80(3): B1–B10.

Damasio, A.R. 1998. Emotion in the perspective of an integrated nervous system. *Brain Research Reviews* 26:83-86.

`

Darwin, C. 1871. *The Descent of Man.* London: Murray.

Davidson, I. 1999. The game of the name: continuity and discontinuity in language origins. In B.J. King (ed) *The origins of language: What nonhuman primates can tell us.* Santa Fe: School of American Research Press, 229-268.

Davidson, R.J. & Erlichman, H. 1980. Lateralized cognitive processes and the electroencephalogram. *Science* 207:1005-1007.

Davies, J.B., Dillon, T., MacDonald, R.A.R., Mitchell, L.A. & Serpell, M.G. 2003. An empirical investigation of the anxiolytic and pain reducing effects of music. *Psychology of Music* 31(2):187-203.

Deacon, T. 1992. The neural circuitry underlying primate calls and human language. In J. Wind, B. Chiarelli, B. Bichakjian, & A. Nocentini (eds) *Language origin: A multidisciplinary approach.* Boston: Kluwer Academic, 121-162.

Deacon, T.W. 1997. *The symbolic species. The co-evolution of language and the human brain.* London: The Penguin Press.

Deacon, T.W. 2000. Evolutionary perspectives on language and brain plasticity. *Journal of Communication Disorders* 33:273-291.

Deacon, T.W. 2003. Universal grammar and semiotic constraints. In M.H. Christiansen & S. Kirby *Language Evolution*. Oxford: Oxford University Press, 111-140.

De Casper, A. J. & Fifer, W. P. 1980. Of human bonding: Newborns prefer their mothers' voices. *Science* 208: 1174–1176.

De l'Etoile, S.K. 2006. Infant behavioral responses to infant-directed singing and other maternal interactions. *Infant Behavior and Development* 29:456-470.

Demaree, H.A., Everhart, D.E., Youngstrom, E.A. & Harrison, D.W. 2005. Brain lateralization of emotional processing: historical roots and a future incorporating "dominance". *Behavioral Cognitive Neuroscience Review* 4(1):3-20.

`

De Paula, R. and Dourish, P. 2005. Cognitive and cultural view of emotions. *Proceedings of the Human Computer Interaction Consortium Winter Meeting* (Douglas, CO).

Dissanayake, E. 2000. Antecedents of the temporal arts in early mother-infant interaction. In N.L. Wallin, B. Merker & S. Brown, S. (eds) *The origins of music.* Cambridge MS: MIT Press, 389-410.

Dissanayake, E. 2001. Becoming Homo Aestheticus: Sources of Aesthetic Imagination in Mother-Infant Interactions *SubStance, Special Issue: On the Origin of Fictions: Interdisciplinary Perspectives:* 30 (94/95): 85-103.

Dissanayake, E. 2005. Review of Mithen, S. 2005. The Singing Neanderthals: the Origins of Music, Language, Mind and Body. London: Weidenfeld & Nicholson. *Evolutionary Psychology* 3:375-380.

Donald, M. 1991. *Origins of the Modern Mind: Three stages in the evolution of culture and cognition.* Cambridge MA: Harvard University Press.

Drake, C. & Bertrand, D. 2001. The quest for universals in temporal processing in music. *Annals of the New York Academy of Sciences* 930:17-27.

Duffy, D. & Ball, G. 2002. Song predicts immunocompetence in male European starlings (*Sturnus vulgaris*). *Proceedings of the Royal Society London.* 01PB0839: 1–6.

Dunbar, R. 1998. Theory of mind and the evolution of language. In: J.R. Hurford, M. Studdert-Kennedy & C. Knight (eds) *Approaches to the evolution of language.* Cambridge: Cambridge University Press, 92-110.

Eck, D., Gasser, M. & Port, R. 2000. Dynamics and embodiment in beat induction. In P. Desain & L. Windson (eds) *Rhythm perception and production.* Lisse: Swets & Zeitlinger, 157-170.

Edelman, G.M. 1992. *Bright Air, Brilliant Fire: On the matter of the Mind.* New York: Basic Books.

Edelman, G.M. 2003. Naturalizing consciousness: a theoretical framework. *Proceedings of the National Academy of Sciences* 100(9): 5520-5524.

`

Ejiri, K. & Masataka, N. 2001. Co-occurrence of preverbal vocal behavior and motor action in early infancy. *Developmental Science* 4(1):40-48.

Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P. & Pääbo, S. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418:869-872.

Falk, D. 2004. Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences* 27:491-541.

Feld, S. & Fox. A.A. 1994. Music and Language. *Annual Review of Anthropology* 23:25-53.

Fernald, A. 1992. Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective In: J. Barkow, L. Cosmides & J. Tooby (eds) *The adapted Mind: Evolutionary psychology and the generation of culture.* Oxford : Oxford University Press, 391-428.

Fisher, S.E. 2006. Tangled webs: tracing the connections between genes and cognition. *Cognition* 101:270-297.

Fishman, Y.I., Volkov, I.O., Noh, M.D., Garell, P.C., Bakken, H., Arezzo, J.C., Howard, M.A. & Steinschider, M. 2001. Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology* 86:2761-2788.

Fitch, W.T. 2000. The evolution of speech: a comparative review. *Trends in Cognitive Sciences* 4(7): 258-267.

Fitch, W.T. 2005a. The evolution of music in comparative perspective. *Annals of the New York Academy of Sciences* 1060:29-49.

Fitch, W.T. 2005b. Dancing to Darwin's tune. Review of Mithen, S. The singing Neanderthals: the Origins of Music, Language, Mind and Body. London: Weidenfeld & Nicholson. *Nature* 238: 488.

Fitch, W.T. 2006a. The biology and evolution of music: A comparative perspective. *Cognition* 100(1): 173-215.

`

Fitch, W.T. 2006b. On the biology and evolution of music. *Music Perception* 24(1): 85-88.

Fraisse, P. 1982. Rhythm and tempo. In D. Deutsch (ed) *The psychology of Music*. New York: Academic Press, 149-180.

Frayer, D.W. & Nicolay, C. 2000. Fossil evidence for the origin of speech sounds In: N.L. Wallin, B. Merker & S. Brown. (eds) *The origins of music*. Cambridge MA: MIT Press, 217-234.

Freeman, W. J. 1995. *Societies of Brains: A Study in the Neuroscience of Love and Hate*. Hillsdale, NJ: Lawrence Erlbaum.

Freeman, W. J. 1999. *How Brains Make Up Their Minds*. London: Weidenfeld and Nicholson.

Freeman, W. J. 2000a. A Neurobiological Role of Music in Social Bonding. In: N. L. Wallin, B. Merker & S. Brown (eds) *The Origins of Music*. Cambridge MA: MIT Press, 411-424.

Freeman, W. J. 2000b. *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics*. London: Springer-Verlag.

Fukui, H. & Yamashita, M. 2003. The effects of music and visual stress on testosterone and cortisol in men and women. *Neuroendocrinology Letters* 24 (3/4): 173-180.

Gamble, C. 2006. Musical Chairs: Review feature: Mithen, S. 2005. The Singing Neanderthals: the Origins of Music, Language, Mind and Body. London: Weidenfeld & Nicholson. *Cambridge Archaeological Journal* 16(1):108-109.

Gandour, J., Wong, D., Dzemitzik, M., Lowe, M., Tong, Y. & Li, X. 2003. A cross-linguistic fMRI study of perception of intonation and emotion in Chinese. *Human Brain Mapping* 18:149-157.

Gannon, P.J. & Laitman, J.T. 1993. Can we see language areas on hominid brain endocasts? *American Yearbook of Physical Anthropology (supplement)* 16:91.

Gannon, P.J., Holloway, R.L., Broadfield, D.C. & Braun, A.R. 1998. Assymetry of chimpanzee planum temporale: humanlike brain pattern of Wernicke's language area homolog. *Science* 279: 222-226.

`

Geissman, T. 2000. Gibbon songs and human music in an evolutionary perspective. In: N.L. Wallin, B. Merker & Brown, S. (eds) *The Origins of Music.* London: MIT Press, 103-123.

Gibson, K.R. 1996. The biocultural human brain, seasonal migrations, and the emergence of the Upper Palaeolithic. In: P. Mellars. & K. Gibson (eds) *Modelling the early human mind*. Cambridge: McDonald Institute Monographs, 33-48.

Gibson, E. J. & Pick, A. D. 2000. *An ecological approach to perceptual learning and development*. Oxford: Oxford University Press.

Glendenning, K.K. 2005. The evolution of auditory pathways underlying human hearing. *The Mankind Quarterly* XLVI: 1-27.

Gogtay, N, Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, C., Nugent III, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L. & Thompson, P.M. 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences* 101(21): 8174-8179.

Goldstein, A. 1980. Thrills in response to music and other stimuli. *Physiological Psychology* 8:126-129.

Gomez, P. & Danuser, B. 2004. Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology* 53: 91-103.

Goodman, C.S. & Coughlin, B.C. 2000. The evolution of evo-devo biology. *Proceedings of the National Academy of Sciences* 97: 4424-25.

Granot, R.Y., Gritsenko, V., Lerer, E., Gritsenko, I., Bachner-Melman, R., Israel, S. & Ebstein, R.P. 2004. Provisional evidence that the arginine vasopressin 1a receptor gene is associated with musical memory. *Evolution and Human Behaviour* 28:313-318.

Gratier, M. 2000. Expressions of belonging: the effect of acculturation on the rhythm and harmony of mother-infant vocal interaction. *Musicae Scientiae* Special Issue, 1999/2000: 93-122.

Grattan, M.P., De Vos, E., Levy, J. & McClintock, M. 1994. Do baby boys naturally lead with the left foot? Research on the asymmetries of movement paterns in newborns and infants. *Zero to Three* August/September : 28-31.

`

Greenberg, G., Partridge, T., Weiss, E. & Haraway, M.M. 1999. Integrative levels, the brain, and the emergence of complex behavior. *Review of General Psychology* 3(3):168-187.

Grewe, O., Nagel, F. Kopiez, R. & Altenmüller, E.   2007. Listening to music as a re-creative process: physiological, psychological, and psychoaccoustical correlates of chills and strong emotions. *Music Perception* 24(3): 297-314.

Groome, L. J., Mooney, D. M., Holland S. B., Smith, Y. B., Atterbury, J. L., & Dykman, R. A. 2000. Temporal pattern and spectral complexity as stimulus parameters for eliciting a cardiac orienting reflex in human fetuses. *Perception & Psychophysics* 62: 313–320.

Groves, C.P. 1989. *A theory of human and primate evolution.* Oxford: Clarendon Press.

Haas, H.S. & Schauenstein, K. 1997. Neuroimmunomodulation via limbic structures - neuroanatomy of psychoimmmunology. *Progress in Neurobiology* 51:195-222.

Hanna, J.L. 1987. *To dance is human – a theory of nonverbal communication.* Chicago: The University of Chicago Press (2nd edition).

Hannon, E. E., & Trehub, S. E. 2005a. Metrical categories in infancy and adulthood. *Psychological Science* 16: 48–55.

Hannon, E. E. & Trehub, S. E. 2005b. Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences* 102:12639–12643

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

Herlenius, E. & Lagercrantz, H. 2001. Neurotransmitters and neuromodulators during early human development. *Early human development* 65: 21-37.

Huron, D. 2001. Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences* 930:43-61.

Izumi, A. 2000. Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustic Society of America* 108: 3073–3078.

`

Janata, P. & Grafton, S.T. 2003. Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience* 6(7): 682-687.

Jeffery, N. & Spoor, F. 2004. Prenatal growth and development of the modern human labyrinth. *Journal of Anatomy* 204(2): 71-92.

Joseph, R. 2000. Fetal brain behavior and cognitive development. *Developmental Review 20: 81-98.*

Jürgens, U. 1992. On the neurobiology of vocal communication. In: H. Papoušek , U. Jürgens & M. Papoušek (eds) *Nonverbal vocal communication: Comparative and Developmental Approaches*. Cambridge: Cambridge University Press, 31-42.

Juslin, P.N. 2001. Communicating emotion in music performance: a review and theoretical framework. In P. Juslin & J. Sloboda (eds) 2001. *Music and emotion: theory and research*, Oxford: Oxford University Press, 309-337.

Juslin, P. & Sloboda, J.   2001. *Music and emotion: theory and research*. Oxford: Oxford University Press.

Kabuto, M., Kageyama, T., Nitta, H., 1993. EEG power spectrum changes due to listening to pleasant music and their relation to relaxation effects. *Nippon Eiseigaku Zasshi* 48: 807–818.

Kappas, A. 2002. The science of emotion as a multidisciplinary research paradigm. *Behavioural Processes* 60:85-98.

Kelley, A.E. & Berridge, K.C. 2002. The neuroscience of natural rewards: relevance to addictive drugs. *Journal of Neuroscience* 22 (9): 3306–3311.

Keverne, E.B. & Curley, J.P. 2008. Epigenetics, brain evolution and behaviour. *Frontiers in neuroendocrinology* 29:3F89-412.

Khalfa, S., Bella, S.D., Roy, M., Peretz, I. & Lupien, S.J. 2003. Effects of Relaxing Music on Salivary Cortisol Level after Psychological Stress. *NewYork Academy of Sciences* 999**:** 374–376 .

`

Kisilevsky, B.S., Hains, S.M.J., Lee, K. Xie, X. Huang, H., Ye, H.H., Zhang, K. & Wang, Z. 2003. Effects of experience of fetal voice recognition. *Psychological Science* 14(3): 220-224.

Kisilevsky, B.S., Hains, S.M.J., Jacquet, A.-Y, Granier-Deferre, C. & Lecanuet, J.P. 2004. Maturation of fetal responses to music. *Developmental Science* 7(5): 550-559.

Kleinginna, P.R. & Kleinginna, A.M. 1981. A categorized list of emotion definitions with suggestions for a consensual definition. *Motivational Emotion* 5: 345-844.

Knyazev, G.G. 2007. Motivation, emotion and their inhibitory control mirrored in brain oscillations. *Neuroscience and Biobehavioural Reviews* 31: 377-395.

Kreutz, G., Bongard, S., Rohrmann, S., Hdapp, V. & Grebe, D. 2004. Effects of choir singing or listening on secretory immunoglobulin A, cortisol and emotional state. *Journal of Behavioural Medicine* 27(6):623-635.

Krumhansl, C.L. 1997. An exploratory study of musical emotions and psycophysiology. *Canadian Journal of Experimental Psychology* 51:336-52.

Krumhansl, C.L. 2000. Rhythm and pitch in music cognition. *Psychological Bulletin* 126:159-179.

Krumhansl, C.L. 2002. Music: A Link Between Cognition and Emotion. *Current Directions in Psychological Science* 11(2): 45-50.

Krumhansl, C.L. & Iverson, P. 1992. Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance* 18 (3): 739-751.

Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. 2001. A novel forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413: 519–523.

Lai, C. S. L., Gerrelli, D., Monaco, A. P., Fisher, S. E., & Copp, A. J. 2003. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* 126: 2455–2462.

`

Lamont, A.M . 2005. What do monkeys' music choices mean? *Trends in Cognitive Sciences* 9(8): 359-361.

Large, E.W. 2000. On synchronizing movements to music. *Human Movement Sciences* 19:527-566.

Large, E.W. & Palmer, C. 2002. Perceiving temporal regularity in music. *Cognitive Science* 26: 527-566.

Lavy, M.M. 2001. *Emotion and the expression of listening to music, a framework for empirical research.* Unpublished D. Phil thesis, University of Cambridge, Cambridge.

Lecuanet, J.P. 1996. Prenatal auditory experiences. In: I. Deliège & J.A. Sloboda (eds) *Musical Beginnings: The Origins and Development of Musical Competence.* New York: Oxford University Press, 3-36.

Le Doux, J. 1996. *The emotional brain*. New York: Simon & Shuster.

Le Doux, J. 2002. *Synaptic Self, how our brains become who we are*. London: Macmillan.

Levitin, D.J. 2006. *This is your brain on Music*. *The Science of a Human Obsession.* New York: Dutton.

Levitin, D.J. & Menon, V. 2003. Musical structure is processed in the "language" areas of the brain: a possible role for Brodmann area 47 in temporal coherence. *Neuroimage* 20: 2142-52.

Lewkowicz, D.J. 2003. Learning and discrimination of audiovisual events in human infants: the hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues. *Developmental Psychology* 39(5):795-804.

Lieberman, D.E. & McCarthy, R.C. 1999. The ontogeny of cranial base angualation in humans and chimpanzees and its implications for reconstructing pharyngeal dimensions. *Journal of Human Evolution* 36:487-517.

Lieberman, P. 2007. The evolution of human speech, its anatomical and neural bases. *Current Anthropology* 48(1):39-66.

`

Longhi, E. & Karmiloff- Smith, A. 2004.   In the beginning was the song: The complex multimodal timing of mother-infant musical interaction Commentary on Falk 2004. *Behavioural and Brain Sciences* 27:516-517.

Longhi, E. 2003. *The temporal structure of mother-infant interactions in musical contexts.* Doctoral dissertation, The University Of Edinburgh.

MacLarnon, A. 1993. The vertebral canal. In: A. Walker & R. Leaky (eds) *The Nariolotome Homo erectus Skeleton.* Cambridge: Harvard University Press, 359-390.

MacLarnon, A. & Hewitt, G. 2004. Increased Breathing Control: another factor in the evolution of human language. *Evolutionary Anthropology* 13:181-197.

MacNeilage, P. F. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21(4): 499–546.

Malloch, S.N. 1999. Mothers and infants and communicative musicality. *Musicae Scientiae* Special Issue, 1999/2000: 59-74.

Marcus, G.F., Vouloumanos, A. & Sag, I.A. 2003. Does Broca's play by the rules? *Natural Neuroscience* 6:651-52.

Marler, P. 2000. Origins of music and speech: Insights from animals. In N.L.Wallin, B. Merker & S. Brown (eds) *The origins of music*. Cambridge MS: MIT Press, 31-48.

Marshall, J.C. & Fink, G.R. 2003. Cerebral localization, then and now. *Neuroimage* 20:S2-S7.

Martinez, I., Rosa, M., Arsuaga, J.-L., Jarabo, P., Quam, R., Lorenzo, C., Gracia, A., Carretero, J.-M., Bermúdez de Castro, J.-M. & Carbonell, E. 2004. Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences* 101:9976-81.

Masataka, N. 1999. Preference for infant-directed singing in 2-day old hearing infants of deaf parents. *Developmental Psychology* 35:1001-1005.

Masataka, N. 2007. Music, evolution and language. *Developmental Science* 10(1): 35-39.

`

Mastropieri, D. & Turkewitz, G. 1999. Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology* 35: 204-214.

McAuley, J.D., Jones, M.R., Holub, S., Johnston, H.M. & Miller, N.S. 2006. The time of our lives: life span development of timing and event tracking. *Journal of Experimental Psychology: General* 135(3):348-367.

McCraty, R., Atkinson, M. & Rein, G. 1996. Music enhances the effect of positive emotional states on salivary IgA. *Stress Medicine* 12:167-175.

McCune, L., Vihman, M.M., Roug-Hellichius, L., Delery, D.B. & Gogate, D.B. 1996. Grunt Communication in Human Infants *(Homo sapiens). Journal of Comparative Psychology* 110 (1):27-37.

McDermott, J. 2008. The evolution of music. *Nature* 453:287-288.

McDermott, J. & Hauser, M. 2004. Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition* 94:B11-B21.

McDermott, J. & Hauser, M. 2007. Nonhuman primates prefer slow tempos but dislike music overall. *Cognition* 104:654-668.

McFarland, R.A. & Kennison, R. 1989. Asymmetry in the relationship between finger temperature changes and emotional state in males. *Biofeedback and Self Regulation* 14:281-290.

McNeill, W. H. 1995. *Keeping Together in Time*. London: Harvard University Press.

Menon, V. & Levitin, D.J. 2005. The reward of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage* 28:175-184.

Merker, B. 2000. Synchronus chorusing and human origins. In N.L.Wallin, B. Merker & S. Brown, S (eds) *The origins of music.* Cambridge MS: MIT Press, 315-327.

Merker, B. 1999. Synchronus chorusing and the origins of music. *Musicae Scientiae.* Special Issue 1999-2000:59-74.

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

`

Miall, D.S. & Dissanayake, E. 2003. The poetics of babytalk. *Human Nature* 14(4):337-364.

Miller, G. 2000. Evolution of human music through sexual selection. In N.L. Wallin, B. Merker & S. Brown (eds) *The origins of music*. Cambridge, MA: MIT Press, 329-360.

Miller, G.F. 2001. *The mating mind.* New York: Doubleday.

Mitani, J., Hunley, K. & Murdoch, M. 1999. Geographic variation in the calls of wild chimpanzees: a reassessment. *American Journal of Primatology* 47: 133–151.

Mitchell, R.L.C., Elliott, R., Barry, M., Cruttende, A. & Woodruff, P.W.R. 2003. The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia* 41:1410-1421.

Mithen, S. 1996. *The prehistory of the mind: a search for the origins of art, religion and science.* London: Thames and Hudson.

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

Mithen, S. 2006. Review Feature: The Singing Neanderthals: the Origins of Music, Language, Mind and Body. *Cambridge Archaeological Journal* 16(1):97-112.

Moggi-Cecchi, J. and Collard, M.   2002. A fossil stapes from Sterkfontein, South Africa, and the hearing capabilities of early hominids. *Journal of Human Evolution* 42(3):259-65.

Monnot, M., Foley, R. & Ross, E. 2004. Affective prosody: whence motherese. Commentary on Falk, D. Prelinguistic evolution in early hominins: whence motherese? *Behavioral and Brain Sciences* 27:519-520.

Morley, I. 2002. Evolution of the physiological and neurological capacities for music. *Cambridge Archaeological Journal* 12(2): 195-216.

Morley, I. 2003. *The evolutionary origins and archaeology of music: an investigation into the prehistory of human musical capacities and behaviours.* PhD thesis, University of Cambridge. [http://ww.dar.cam.ac.uk/dcrr.]

`

Morley, I. 2006. Name that tune. Review feature: Mithen, S. 2005. The Singing Neanderthals: the Origins of Music, Language, Mind and Body. London: Weidenfeld & Nicholson. *Cambridge Archaeological Journal* 16(1):101-103.

Müller, R.A. 1996. Innateness, autonomy, universality? Neurobiological approaches to language. *Behavioral and Brain Sciences* 19:611-675.

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M. & Matsuzawa, T. 2004. Imitation in neonatal chimpanzees (*Pan troglodytes*). *Developmental Science* 7(4):437–442.

Nakata, T. & Trehub, S. E. 2004. Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development* 27: 455–464.

Newman, J.D. 2004. Motherese by any other name: Mother-infant communication in non-hominin mammals. Commentary on Falk 2004. *Behavioural and Brain Sciences* 27: 519-520.

Nicholson, K.G., Baum, S., Kilgour, A., Koh, C.K., Munhall, G. & Cuddy, L.L. 2003. Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain & Cognition* 52:382-389.

Nishimura, T. 2003. Comparative morphology of the hyo-laryngeal complex in anthropoids: two steps in the evolution of the larynx. *Primates* 44: 41-49.

Nishimura, T. 2005. Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. *American Journal of Physical Anthropology* 126:193-204.

Nishimura, T. 2006. Descent of the larynx in chimpanzees: mosaic and multiple-step evolution of the foundations for human speech. In: T. Matsuzawa, M. Tomonaga & M. Tanaka (eds) *Cognitive Development in Chimpanzees*. Tokyo:Springer-Verlag, 75-95.

Nishimura, T., Mikami, A, Suzuki, J & Matszawa, T. 2003. Descent of the larynx in chimpanzee infants. *Proceedings of the National Academy of Sciences* 100:6930-6933.

Noble, W. & Davidson, I. 1997. Reply to Mithen. *Cambridge Archaeological Journal* 7: 279-284.

Panksepp, J. 1995. The emotional sources of "chills" induced by music. *Music Perception* 13:171-207.

`

Panksepp, J. 1998a. *Affective neuroscience: the foundations of human and animal emotions*. Oxford University Press: New York.

Panksepp, J. 1998b. The periconscious substrates of consciousness: Affective states and the evolutionary origins of the self. *Journal of Consciousness Studies* 5: 566-582.

Panksepp, J. 2001. The neuro-evolutionary cusp between emotions and cognition. *Evolution and Cognition* 7(2):141-163.

Panksepp, J. & Panksepp, J.B. 2000. The seven sins of evolutionay psychology. *Evolution and Cognition* 6(2):108-131.

Panksepp, J. & Bernatzky, G. 2002. Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. *Behavioural Processes* 60:133-155.

Papaeliou, C. & Trevarthan, C.   1998. The infancy of music. *Musical Praxis* 1(2):19-33.

Papaeliou, C., Minadakis, G. & Cavouras, D. 2002. Acoustic Patterns of Infant Vocalizations Expressing emotions and Communicative Functions. *Journal of Speech, Language and Hearing Research* 45:311-317.

Papoušek, H. 1996. Musicality in infancy research. In: I. Deliège & J.A. Sloboda (eds) *Musical Beginnings.* New York: Oxford University Press, 37-55.

Papoušek, M. 1996. Intuitive parenting: a hidden source of musical stimulation in infancy In: I. Deliège & J.A. Sloboda (eds) *Musical Beginnings.* New York: Oxford University Press, 88-112.

Papoušek, M., Bornsgein, M.H., Nuzzo, C., Papoušek, H. & Symmes, D. 1990. Infant responses to prototypical melodic contours in parental speech. *Infant behaviour and development* 13: 539-45.

Papoušek , M., Papoušek , H. & Symmes, D. 1991. The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development* 14:415-440.

Parncutt, R. *in press a*. Prenatal development. In S. Hallam, I. Cross. & M. Thaut (eds) *Oxford handbook of music psychology.* Oxford: Oxford University Press.

`

Parncutt, R. *in press b.* Prenatal and infant conditioning, the mother schema, and the origins of music and religion. *Musicae Scientiae, in press. Special issue on music and evolution edited by Oliver Vitouch and Olivia Ladinig.*

Parr, L.A. & Hopkins, W.D. 2000. Brain temperature asymmetries and emotional perception in chimpanzees, *Pan troglodytes. Physiology and Behaviour* 71:363-371.

Parsons, L. 2003. Exploring the functional neuroanatomy of music performance, perception and comprehension. In L. Peretz. & R. Zatorre (eds) *The Cognitive Neuroscience of Music*, Oxford: Oxford University Press, 247-68.

Patel, A.D. 2003. Language, music, syntax and the brain. *Nature Neuroscience* 6:674-81.

Patel, A.D. 2006. Musical rhythm, linguistic rhythm and human evolution. *Music Perception* 24(1): 99-104.

Patel, A.D. & Balaban, E. 2000. Temporal patterns of human cortical activity reflect tone sequence structure. *Nature* 404:80–84.

Patel, A.D. & Daniele, J.R. 2003. An empirical comparison of rhythm in language and music. *Cognition* 87:B35-B45.

Patel, A.D., Peretz, I., Tramo, M. & Labreque, R. 1998. Processing prosodic and music patterns: a neuropsychological investigation. *Brain and Language* 61:123-44.

Paterson, S.J., Heim, S., Friedman, J.T., Coudhury, N. & Benasich, A.A. 2006. Development of structure and function in the infant brain: implications for cognition, language and social behaviour. *Neuroscience and Biobehavioral Reviews* 30:1087-1105.

Peretz, I. 2001. Listen to the brain, the biological perspective on musical emotions. In P. Juslin & J. Sloboda (eds) *Music and emotion: theory and research* Oxford: Oxford University Press, 105-134.

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

Peretz, I. & Gagnon, L. 1999. Dissociation between recognition and emotional judgment for melodies. *Neurocase* 5:21-30.

`

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

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

Peretz, I., Gagnon, L. & Bouchard, B. 1998. Music and emotion: perceptual determinants, immediacy and isolation after brain damage. *Cognition* 68:111-141.

Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E. & Evans, A.C. 1999. Localization of cerebral activity during simple singing. *NeuroReport* 10:3979-3984.

Petitto, L. A. & Marentette, P. F. 1991. Babbling in the manual mode: evidence for the ontogeny of language. *Science* 251: 1483–1496.

Petitto, L.A., Holowka, S., Sergio, L.E., Levy, B. & Ostry, D.J. 2004. Baby hands that move to the rhythm of language: hearing babies acquiring sign languages babble silently on the hands. *Cognition* 93:43-73.

Phillips-Silver, J. & Trainor, L.J. 2005. Feeling the beat: movement influences infant rhythm perception. *Science* 308: 1430.

Phillips-Silver, J. & Trainor, L. J. 2007. Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition* 105: 533–546.

Phillips-Silver, J. & Trainor, L. J. 2008. Vestibular influene on auditory metrical interpretation. *Brain & Cognition* 67: 94-102.

Pierce, J.R. 1996. *The Science of Musical Sound*. San Fransisco:W.H. Freeman (3rd edition)

Pinker, S. 1997. *How the mind works.* New York: W.W. Norton.

Pouthas, V. 1996. The development of the perception of time and temporal regulation of action in infants and children. In I. Deliège & J. A. Sloboda (eds) *Musical beginnings: Origins and development of musical competence*. New York: Oxford University Press, 115–141.

Preuss, T.M. 2006. Who's afraid of *Homo sapiens*? *Journal of Biomedical discovery and collaboration* 1:1-12.

`

Previc F.H. 1999. Dopamine and the origins of human intelligence. *Brain and Cognition* 41:299-350.

Provine, R.R. 2004. Laughing, Tickling, and the Evolution of Speech and Self . *Current Directions in Psychological Science* 13 (6), 215–218.

Rapp, B. 2001. *The handbook of cognitive neuropsychology. What deficits reveal about the human mind.* Philadelphia P.A.: Psychology Press.

Rauschecker, J.P. 2002. Where science meets the arts. Review of Benzon, W.L. 2001. Beethoven's anvil: Music in mind and Culture. *Science* 296:1032.

Reed, S.A. 1998. The politics and poetics of dance. *Annual Review of Anthropology* 27:503-532.

Repp, B. H. 2005. Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review* 12: 969-992.

Richman, B. 1993. On the evolution of speech: Singing as the middle term. *Current Anthropology* 34:721-722.

Richman, B. 2000. How Music Fixed "Nonsense" into Significant Formulas: On Rhythm, Repetition, and Meaning. In: N.L. Wallin, B. Merker & S. Brown (eds) T*he Origins of Music*. London: MIT Press, 301-314.

Richmond, B.G., Aiello, L.C. & Wood, B.A. 2002. Early hominin limb proportions. *Journal of Human Evolution* 43: 529–548.

Richmond, B.G., Jungers, W.L. 2008. *Orrorin tugenensis* femoral morphology and evolution of hominin bipedalism. *Science* 319:1662-1664.

Rickard, N.S. 2004. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music* 32(4):371-388.

Rightmire, G.P. 1998. Human evolution in the Middle Pleistocene: the role of *Homo heidelbergensis. Evolutionary Anthropology* 6:218-227.

Rilling, J.K. 2006. Human and nonuman primate brains: are they allometrically scaled versions of the same design? *Evolutionary Anthropology* 15:65-77.

`

Robb, L. 2000. Emotional musicality in mother-infant vocal affect, an acoustic study of postnatal depression. *Musicae Scientiae* Special Issue 1999/2000:123-154.

Ross, E.D. 2000. Affective prosody and the aprosodias. In: M. Mesulam (ed) *Principles of behavioral and cognitive neurology.* New York: Oxford University Press, 316–331.

Roth, G. & Dicke, U. 2005. Evolution of the brain and intelligence. *Trends in Cognitive Sciences* 19(5): 250-257.

Ruff, C.B., Trinkhaus, E. & Holliday, T.W. 1997. Body mass and encephalization in Pleistocene *Homo*. *Nature* 387:173-176.

Schellenberg, E.G. & Trehub, S.E. 1996. Natural musical intervals: evidence from infant listeners. *Psychological Science* 7**:** 272–277.

Scherer, R.C. 1991. Emotion expression speech and music. In J. Sundberg, L. Nord & R. Carlson (eds) *Music, Language, speech and brain*. MacMillan Press: Basingstoke, 146-156.

Scherer, K.R. 1995. Expression of emotion in voice and music. *Journal of Voice* 9(3):235-248.

Scherer, K.R. 2003. Vocal communication of emotion: a review of research paradigms. *Speech Communication* 40(1-2): 227 - 256 .

Scherer, K.R. 2004. Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research* 33: 239-251.

Schmidt, l.A. & Trainor, L.J. 2001. Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognition and Emotion* 15:487-500.

Schwartz, G.E., Davidson, R.J. & Maier, F. 1975. Right hemisphere lateralization for emotion in the human brain; interactions with cognition. *Science* 190:286-288.

Shenfield, T., Trehub, S.E. & Nakata, T. 2003. Maternal singing modulates infant arousal. *Psychology of Music* 31(4): 365-375.

Shepard, J. & Wicke, P. 1997. *Music and Cultural Theory*. Oxford: Blackwill.

`

Shimai, S. 1992. Emotion and identification of environmental sounds and electroencephalographic activity. *Fukushima Journal of Medical Science* 38: 43–56.

Sidtis, J.J. 2000. From chronograph to functional image: what's next? *Brain and Cognition* 42:75-77.

Skoyles, J.R. 2000. Without breath and without song. *Trends in Cognitive Sciences* 4(11):405.

Sloboda, J.A. 1991. Music structure and emotional response: some empirical findings. *Psychology of Music* 19:110-120.

Small, C. 1998. *Musicking: The meanings of performing and listening*. Hanover, NH: Wesleyan University Press.

Smith, L.B. 2006. Movement matters: the contributions of Esther Thelen. *Biological Theory* 1(1):87-89.

Smith, L. & Gasson, M. 2005. The development of embodied cognition: six lessons from babies. *Artificial Life* 11(1-2):13-29.

Snow, D. 2000. The emotional basis of linguistic and nonlinguistic intonation: implications for hemispheric specialization. *Developmental Neuropsychology* 17:1-28.

Soressi, M. 2005. Late Mousterian lithic technology: its implications for the pace of emergence of behavioural modernity and the relationship between behavioural modernity and biological modernity. Proceedings of the International Round Table Conference: *From tools to symbols- from early hominids to modern humans*. WITS: South Africa, 389-417.

Spoor, C.F., Wood, B.A. & Zonneveld, F. 1994. Evidence for a link between the human semicircular canal size and bipedal behavior. *Journal of Human Evolution* 30:183-187.

Spoor, F., Hublin, J., Braunc, M. & Zonneveld, F. 2003. The bony labyrinth of Neanderthals. *Journal of Human Evolution* 44: 141–165.

Spoor, F., Garland, T., Krovitz, G., Ryan, T. M., Silcox, M.T. & Walker, A. 2007. The primate semicircular canal system and locomotion. *Proceedings of the National Academy of Sciences* 104(26): 10808–10812.

`

Stewart, L. & Walsh, V. 2005. Infant Learning: music and the baby brain. *Current Biology* 15(21):R882-R884.

Stokstad, E. 2003. Peering into ancient ears. *Science* 302: 770–771.

Stow, G.W. 1905. *The native races of South Africa*. London:Swan Sonnenschein.

Strogatz, S. H. & Stewart, I. 1993. Coupled Oscillators and Biological Synchronization. *Scientific American* December: 102-109.

Sundberg, J. 1987. *The science of the singing voice*. Dekalb, Illinois: Northern Illinois University Press.

Tallerman, M. 2006. Abracadabra! Early hominin for 'I think my humming's out of tune with the rest of the world!" Review feature: Mithen, S. 2005. The Singing Neanderthals: the Origins of Music, Language, Mind and Body. *Cambridge Archaeological Journal* 16(1):106-107.

Tallerman, M. 2007. Did our ancestors speak a holistic protolanguage? *Lingua*: 579-604.

Thaut, M. H., McIntosh, G. C. & Rice, R. R. 1997. Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation. *Journal of Neurological Sciences* 151: 207-212.

Thelen, E. 1981. Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology* 17:237–257.

Thelen, E. & Smith, L.B. 1998. *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Thompson, F.T., Schellenberg, E.G. & Husain, G. 2004. Decoding speech prosody: do music lessons help? *Emotion* 4(1):46-64.

Thomson, W. 2004. From sounds to music: the contextualizations of pitch. *Music Perception* 21(3):431-456.

Titze I. 1995. Motor and sensory components of a feedback-control model of fundamental frequency. In: F. Bell-Berti & L. Kaphael (eds) *Producing Speech: Contemporary Issues* New York: American Institute of Physics, 309-318.

`

Toates, F. 2001. *Biological Psycology: An Integrative Approach.* Harlow: Pearson Education.

Todd, N.P. 1999. Motion in music: a neurobiological perspective. *Music Perception* 17:115-126.

Todd, N.P. & Lee, C.S. 2007. Reply to "Embodied Rhythm" by Bruno Repp and "Do Preferred Beat Rate and Entrainment to the Beat Have a Common Origin in Movement?" by Laurel Trainor. *Empirical Musicology Review* 2(3):110-112.

Todd, N. P., Cousins, R. & Lee, C. S. 2007. The contribution of anthropomorphic factors to individual differences in the perception of rhythm. *Empirical Musicology Review* 2(1): 1–13.

Toga, A.W. & Thompson, P.M. 2003. Mapping brain asymmetry. *Nature Reviews Neuroscience* 5:37-47.

Tolbert, E. 2001. Music and meaning: an evolutionary story. *Psychology of Music* 29:84-94.

Trainor, L. 2007. Do preferred beat rate and entrainment to the beat have a common origin in movement? *Empirical Musicology Review* 2(1): 17-20.

Trainor, L. 2008. The neural roots of music. *Nature* 453:598-599.

Trainor, L.J. & Heinmiller, B.M. 1998. The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development* 21:77-88.

Trainor, L., Clark, E., Huntley, A. & Adams, B. 1997. The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development* 20:383-396.

Trainor, L.J., Austin, C.M. & Desjardins, R.N. 2000. Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science* 11:188-195.

Trainor, L.J., Tsang, C.D. & Cheng, V.H.W. 2002. Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception* 20:187-194.

Tramo, M.J., Cariani, PA., Delgutte, B. & Braida, L.D. 2001. Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Sciences* 930: 92-116.

`

Trehub, S. E. 2001. Musical Predisposition in Infancy. *Annals of the New York Academy of Sciences* 930:1-16.

Trehub, S.E. 2003. The developmental origins of musicality. *Nature Neuroscience* 6(7):669-673.

Trehub, S.E., Trainor, L.J. & Unyk, A.M. 1993. Music and speech processing in the first year of life. In: H.W. Reese (ed) *Advances in Child Development and Behavior*. New York: Academic Press, 24:1-35.

Trevarthen, C. 1998. The concept and foundations of infant intersubjectivity. In S. Braten (ed) *Intersubjective communication and emotion in early ontogeny*. Cambridge: Cambridge University Press, 15–46.

Trevarthen, C. 1999. Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae* Special Issue:155-215.

Trevarthen, C. 2005. First things first: infants make good use of the sympathetic rhythm of imitation, without reason or language. *Journal of Child Psychotherapy* 31(1): 91 – 113.

Trevarthen, C. & Aitken, K.J. 2001. Infant intersubjectivity: Research, theory and clinical applications. *Journal of Child Psychology and Psychiatry* 42(1):3-48.

Trinkhaus E. 1984. Comment on: Carrier, D.R. Human running and hominid evolution. *Current Anthropology* 25:491.

Vallortigara, G., Rogers, L.J. & Bisazza, A. 1999. Possible evolutionary origins of cognitive brain lateralization. *Brain Research Reviews* 30:164-175.

Van Lancker, D. & Sidtis, J. J. 1992. The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: all errors are not created equal. *Journal of Speech and Hearing Research* 35: 963–970.

Vaneechoutte, M. & Skoyles, J.R. 1998. The memetic origin of language: modern humans as musical primates. *Journal of Memetics – Evolutionary Models of Information Transmission 2* (http://jom-emit.cfpm.org/1998/vol2)

`

Viljoen, M., Panzer, A., Roos, J. & Bodemer, W. 2003. Psychoneuroimmunology: from philosophy, intuition and folklore to a recognized science. *South African Journal of Science* 99:332-336.

Vilkman, E., Sonninen, A., Hurme, P. & Körkkö, P. 1996. External Laryngeal Frame Function in Voice Production Revisited: A Review. *Journal of Voice* 10(1):78-92.

Vouloumanos, A. & Werker, J.F. 2007. Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science* 10(2):159-171.

Wallin, N.L. 1991. *Biomusicology.* Stuyvesant, NY: Pendragon Press.

Weaver, A.H. 2005. Reciprocal evolution of the cerebellum and neocortex in fossil humans. *Proceedings of the National Academy of Sciences, USA.* 102:3576-3580.

Wilkens, W.K. & Wakefield, J. 1996. Further issues in neurolinguistic preconditions. *Behavioral and Brain Sciences* 19(4): 793-798.

Williams, L. 1967. *The dancing chimpanzee: a study of primitive music in relation to the vocalizaing and rhythmic action of apes.* New York: Norton.

Winkelman, M. 2002. Shamanism as neurotheology and evolutionary psychology. *American Behavioral Scientist* 45(12):1873-1885.

Wood, B. A. 1992. Origin and evolution of the genus *Homo*. *Nature* 355**:** 783–790.

Wray, A. 1998. Protolanguage as a holistic system for social interaction. *Language and Communication* 18:47-67.

Wray. A. 2000. Holistic utterances in protolanguage: the link from primates to humans. In C. Knight, M. Studdert-Kennedy & J.R. Hurford (eds) *The evolutionary emergence of language: social function and the origins of linguistic form.* Cambridge: Cambridge University Press, 285-302.

Wray, A. 2006. Joining the dots: the evolutionary picture of language and music. Review feature: Mithen, S. 2005. The Singing Neanderthals: the Origins of Music, Language, Mind and Body. *Cambridge Archaeological Journal* 16(1):103-15.

`

Wray, A. & Grace, G.W. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117:543-578.

Xu, Y. 2005. Speech melody as articulatory implemented communicative function. *Speech Communication* 46:220-251.

Zatorre, R.J., Belin, P. & Penhune, V.B. 2002. Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences* 6(1): 37-46.

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

Zentner, M. R. & Kagan, J. 1998. Infants' perception of consonance and dissonance in music. *Infant Behavior and Development* 21: 483–492.

Zilhão, J. 2007. The emergence of ornaments and art: an archaeological perspective on the origins of "behavioural modernity". *Journal of Archaeological Research* 15:1-54.