doi 10.4436/JASS.92008

JASs Invited Reviews
Journal of Anthropological Sciences
Vol. 92 (2014), pp. 147-177

A multi-disciplinary approach to the origins of music: perspectives from anthropology, archaeology, cognition and behaviour

Iain Morley

Institute of Human Sciences, Institute of Cognitive & Evolutionary Anthropology, 58a Banbury Road, Oxford, U.K.
e-mail: iain.morley@anthro.ox.ac.uk

Summary - *Archaeological evidence for musical activities pre-dates even the earliest-known cave art and it remains the case that no human culture has yet been encountered that does not practise some recognisably musical activity. Yet the human abilities to make and appreciate music have been described as "amongst the most mysterious with which [we are] endowed" (Charles Darwin, 1872) and music itself as "the supreme mystery of the science of man" (Claude Levi-Strauss, 1970). Like language, music has been the subject of keen investigation across a great diversity of fields, from neuroscience and psychology, to ethnography, to studies of its structures in its own dedicated field, musicology; unlike the evolution of human language abilities, it is only recently that the origins of musical capacities have begun to receive dedicated attention. It is increasingly clear that human musical abilities are fundamentally related to other important human abilities, yet much remains mysterious about this ubiquitous human phenomenon, not least its prehistoric origins. It is evident that no single field of investigation can address the wide range of issues relevant to answering the question of music's origins. This review brings together evidence from a wide range of anthropological and human sciences, including palaeoanthropology, archaeology, neuroscience, primatology and developmental psychology, in an attempt to elucidate the nature of the foundations of music, how they have evolved, and how they are related to capabilities underlying other important human behaviours. It is proposed that at their most fundamental level musical behaviours (including both vocalisation and dance) are forms of deliberate metrically-organised gesture, and constitute a specialised use of systems dedicated to the expression and comprehension of social and emotional information between individuals. The abilities underlying these behaviours are selectively advantageous themselves; in addition, various mechanisms by which the practice of musical activities themselves could be advantageous are outlined.*

Keywords - *Vocalisation, Sociality, Entrainment, Hominin, Human Evolution, Palaeoanthropology, Palaeolithic, Hunter-gatherer.*

The human capabilities for carrying out musical behaviours have been described as "amongst the most mysterious with which [we are] endowed" (Darwin, 1871) and music itself as "the supreme mystery of the science of man" (Levi-Strauss, 1970). The reasons for this are manifold, but most conspicuous amongst them are music's uniqueness in humans and universality amongst human cultures, whilst, on the surface of it, serving no obvious immediate benefit for survival.

Evolutionary perspectives on human cognition and behaviour have burgeoned in the last 30 years, adding to our understanding of certain aspects of human behaviour which have core common elements underlying the variation that exists across different cultures, by considering human behavioural capabilities and tendencies in light of the biological history of our species, and the selective pressures to which our ancestors were subject. Foci have included morality, supernatural beliefs, theory of mind, aesthetic preferences, language, symbolism and culture, amongst other things (Laland & Brown, 2011, offer a very useful overview and critique of different approaches that have been taken; Barrett *et al.,* 2001 provide useful introduction to the questions and concerns of human evolutionary psychology in particular).

Until the last decade or so music has been underrepresented in such investigations and theorizing, both in terms of seeking to understand music's possible roles in our evolutionary history (and the roles of the underlying abilities that support musical behaviours, which is a different concern), and in terms of understanding the relationships between musical capabilities and those supporting other aspects of human behaviour. Increasingly, however, researchers involved in studying music in various fields have started to incorporate evolutionary thinking into their interpretations of their data, and some authors have sought to situate musical capabilities in broader frameworks of human evolution (see, for example, Wallin *et al.*, 2000; Morley, 2002, 2003, 2013; d'Errico *et al.*, 2003; Mithen, 2005; Conard *et al.*, 2009; Malloch & Trevarthen, 2009b; Bannan, 2012; Schulkin, 2013).

Musical behaviours today are peculiar in being simultaneously viewed by some as a functionally redundant leisure activity that we carry out as a sort of 'luxury' addition to our survival activities, and by others as fundamental to a wide range of aspects of human life. In a sense, aspects of Western societies' consumption of music encompass both these perspectives simultaneously: in Western societies in particular music is packaged and sold as a product, a consumable luxury addition to life's essentials, yet these sales are often predicated on the idea that music deals with fundamental human emotional concerns; it is used to adjust consumer behaviour, solicit our affections and votes, and to elicit particular emotional responses in specific circumstances.

It might be possible to view music as purely a cultural product (and indeed, many authors have argued that it is so – for example, Repp, 1991) were it not for the facts that musical activities are practised by all known human societies and that it is increasingly clear that we have several innate and finely-honed cognitive mechanisms that particularly respond to musical stimuli and make their production possible, and appealing. The ubiquity of musical behaviours in human societies requires some explanation, as do the relationships that appear to exist between the capabilities for musical behaviours and for other aspects of human behaviour. Indeed, it is in these relationships that we may find some clues to the evolutionary history of the behaviours. It has become increasingly evident in recent years that the capacities for musical behaviours and those related to complex social behaviours are closely intertwined, and the one set cannot be fully understood without understanding the other. Clearly, musical behaviours are enormously diverse across different societies, and the influence of culture upon the form, role and attributed significance of musical activities is very great. Nevertheless, there are significant commonalities in form and function of musical behaviours, and the forms that these behaviours take are in some important respects shaped and constrained by biological factors.

We know that musical behaviour has a very ancient pedigree. Direct evidence, in the form of musical instruments, appears in the archaeological record at least 40,000 years ago (Conard *et al.*, 2009; Higham *et al.,* 2012). However, this is some 150,000 years after the estimated emergence of our species, *Homo sapiens*, and studies in developmental psychology and neuroscience strongly suggest that the human capacities that underpin musical production and perception have a much longer evolutionary history.

If we are to investigate the longer-term history of musical behaviours, we need to identify which aspects of musical activities are traits that emerge in particular cultural contexts and which aspects are innate and shared between humans. Of the innate elements we are also concerned to distinguish between those which are shared with our nearest surviving primate relatives and those which are exclusive to our own lineage, having emerged since the divergence of hominins from our last common ancestor with chimpanzees. In the latter case we must then also seek to understand the circumstances under which such capabilities emerged, and why.

Music itself is notoriously difficult to define (a matter discussed further below), being both apparently ubiquitous and yet highly varied. The more closely we look at musical behaviours the more it becomes clear that, whilst particular areas of study individually add much to our understanding of the phenomenon and the capacities which make it possible, a full understanding of musicality and its place in our evolution cannot be attempted without drawing upon a very wide range of investigative disciplines, and considering their findings in light of each other. Because of the range of physical and mental capabilities that are used in musical activity, and because of the diversity of ways in which these activities are related to other aspects of behaviour in different human societies, understanding the cross-cultural human phenomenon of music from an evolutionary perspective has to make recourse to many different types of evidence. This means utilising all three of the traditional major anthropological disciplines – biological anthropology (including palaeoanthropology), cultural anthropology (including ethnomusicology) and archaeology. It also means drawing upon other human sciences in the broader sense, including neurological studies, cognitive and developmental psychology, and evolutionary theory, applied to both biology and behaviour. We must also look further afield to the capabilities and behaviours of other animals, especially the other higher primates.

The following discussion highlights some of the relevant findings from this range of disciplines, with the aim of drawing conclusions from the mutual implications of the different evidence. It necessarily constitutes a highly condensed synthesis of complex findings from a wide range of different types of studies; those interested in more detailed examination of these findings and their implications are referred to Morley (2013).

This paper starts by attempting to identify what the principal elements of the investigation must be and considering how we can seek to understand them in an evolutionary context. This is followed by a brief outline of the earliest direct evidence for recognisable musical behaviours in the archaeological record, and of the palaeoanthropological evidence for the evolution of the physiological and neurological features used in musical behaviour. We then look at relationships between vocal tonal abilities, speech, and other forms of vocal communication in higher primates and human infants, to gain some insights into which of these abilities are innate, and the roles that they fulfil in non-linguistic communication. Vocalisation is then also considered in the context of body language and other forms of gestural communication, including the involvement of temporally-coordinated movement and its relationship with emotional experience. Having considered the relationships between these capacities for musicality, and the important roles that they fulfil in human interaction, the discussion finally turns to the question of whether there are ways in which musical activities themselves could confer selective advantages on those carrying them out. Throughout, it will be important to make clear a distinction between selection for the underlying capacities that are used in musical behaviours (the selection for which might or might not be related to the function they fulfil in the musical context), and selection for musical behaviours themselves (i.e. possible selective advantages associated with carrying out musical behaviours).

Conceptualising 'music' and 'musicality' for the purposes of evolutionary analysis

The first difficulty that we face is of defining the focus of our interest, the broad and diverse, but apparently universal entity 'music'. Whilst it is widely asserted in literature dealing with music psychology and anthropology that all human societies carry out behaviours that are recognisable as music (e.g. Clynes, 1982; Storr, 1992; Brown *et al.,* 2000; Blacking, 1995), definitions of exactly what constitutes music very rarely feature in academic papers on the subject. Naturally, the various different disciplines that concern themselves with the study of music have their own preoccupations with particular aspects of music. Thus we find that in studies of music cognition, the focus is usually on the production and processing of particular component parts of music, such as pitches, transitions between them, or tempo. Similarly, developmental studies looking at the emergence of music cognition in human infants are by necessity required to study specific aspects of music perception and production, or of behaviours occurring in relation to pieces of (usually Western) music as whole entities. Neither requires a conception of music as a whole to be defined. Ethnographic studies can describe the form and role of musical activities without the need (or, indeed, the desire) to categorise them under an overarching conceptualization; in the case of the archaeological identification of musical activities, discussion of music is by necessity usually tied to interpretation of instruments themselves and their contexts of use.

Clearly, if we are to examine music through time, across cultures, and across different disciplines, we will need to tackle the difficult issue of describing the focus of our investigations in a way that encompasses the numerous constituent components of those activities, and the diversity of phenomena observed as a whole, but without being so broad as to be meaningless. On the other hand, part of the motivation for the investigation itself is to better understand what music *is* (at the level of the features shared by different societies' behaviours, at least), and how it achieves the effects that it does; the investigation ought to be able to allow us to better formulate a definition than we would be able to otherwise. So we also need to make sure that any conception of music that we start with is not so narrow as to lead us to only investigate what we already think we understand it to be.

There are, of course, dictionary definitions of music, but these tend to describe to a greater or lesser extent music as it exists in the modern Western context, and don't adequately encompass the diversity of forms, effects, uses and conceptions of music that exist across human cultures (see Cross & Morley, 2009 for discussion). Whilst musical behaviours are enormously varied, and different cultures' own conceptions of music are as well, being frequently inseparable from dance and cosmology (e.g. Waterman, 1991; Bohlman, 2002), it remains the case that it is possible to recognise this diverse range of activities across all cultures as 'musical', suggesting that there is an identifiable set of common characteristics, occurring in various combinations. According to Nettl (2000), 'All societies have vocal music. . . . All societies have at least some music that conforms to a meter or contains a pulse. . . . All societies have some music that uses only three or four pitches, usually combining major seconds and minor thirds' (Nettl, 2000, p. 468). There is commonly a division of the octave into five to seven discrete pitches (Burns, 1999; Trehub, 2003), which tend to be separated unequally by tones and semitones. This unequal separation of tonal intervals appears to be a product of the human perceptual system being better able to process unequal scale steps (Butler, 1989; Shepard, 1982), and this feature is innate, occurring from infancy (Trehub *et al.*, 1999). Other products of universal features of the human perceptual system include the ubiquity of the perfect fifth interval, which is more easily processed than other intervals (Schellenberg & Trehub, 1994, 1996a,b), and the perception of consonance, harmony and dissonance (Tramo *et al.*, 2003), with dissonance eliciting aversive reactions from infancy (Trehub, 2003; Gosselin *et al.*, 2006).

It would appear, then, that musical behaviours amongst all humans involve the organisation of sounds into pitches (frequently three to seven), unequally separated across the scale, including the perfect fifth interval, and favouring consonance over dissonance; they involve organising sound sequences so that they have a deliberate structured temporal relationship with each other, including attributing a regular beat to these stimuli (cf. Peretz, 2003).

Musical behaviours also involve intentional bodily action that is temporally organised. Music is impossible to make without action (Besson & Schön, 2003) – it is an embodied activity, not purely an auditory phenomenon, and both rhythmic and tonal sound production unavoidably involve precise, planned, control and sequencing of bodily action. These actions are structured and intentional (Turner & Ioannides, 2009), temporally-organised, and have the potential to have multiple interpretations (Cross, 2003b). It can be seen that, with the exception of the necessity for pitch encoding, the word 'dance' could be substituted for the word 'music' in the preceding sentences. Relationships – and interdependences – between music and dance will be discussed further below.

In summary, on the basis of the above pan-cultural properties of music, and taking into account the broader descriptions offered by the other writers mentioned, musical activities, regardless of whatever other properties and significances they possess in their cultural context, rely on the ability to voluntarily produce sequences of sounds moderated for intensity and/or pitch and/or contour, generated by metrically-organised muscular movements, and often coordinated (entrained) with an internally or externally-perceived pulse. They also comprise the ability to process and extract information from such sounds. Dance clearly also involves voluntary generation of finely-controlled metrically-organised muscular movements, coordinated with an internally or externally-perceived pulse. The investigation of the prehistory of musical behaviours must thus be concerned with the prehistory and evolution of these abilities, their effects, and their relationships with each other and with other human abilities.

How can we study the evolution of music?

Musical stimuli themselves are obviously transitory auditory (and bodily) phenomena that do not preserve over time, so investigating music's prehistory clearly includes challenges. What we can seek to investigate is the prehistory of the capacities that make musical behaviours possible, how they relate to each other and to other human behaviours. We can also investigate the earliest direct evidence for musical behaviours themselves and, in light of what we know about the capacities that make such behaviours possible and what we know about the variations and commonalities in musical behaviours in various cultures, interpret that evidence.

In order to understand the evolutionary history of the capacities that support musical behaviours we need to identify which elements of musical behaviours are innate, being part of our biological heritage. Of each of the innate elements, we need to understand why it is innate, why it is significant, and how it is related to music. We also need to identify how far back in that heritage they go. Which of the innate elements are possessed by other higher primates too, and are thus part of the longer evolutionary history of our species, but providing foundations for the human-specific behaviours that have emerged since? What roles do they fulfil in the context of the higher primates' usage? Which of the innate elements, in contrast, are possessed by humans alone, and have thus apparently emerged in our evolutionary lineage since our last common ancestor with chimpanzees?

In each case an essential element of understanding the history of these capacities is to understand how they are related to other abilities that exist within our primate lineage, such as vocal communication, body language, gesture, imitation, coordination and maintaining complex social relationships, and those that are apparently exclusive to humans, such as linguistic communication and systematic use of symbolism. In addition we can also seek to track the development of some of them in the fossil record of hominin physiological evolution.

The earliest direct evidence of musical activity: Palaeolithic music archaeology

We know that some of the earliest *H. sapiens* in Europe were manufacturing technologically sophisticated musical instruments at least 40,000 years ago (Conard *et al.*, 2009; Higham *et al.* 2012). Bird bone and mammoth ivory pipes (or 'flutes') are found in Aurignacian contexts (those associated with early *H. sapiens* populations) at sites in the Ach and Lone Valleys, Germany (Geissenklösterle, Hohle Fels and Vogelherd) (Hahn & Münzel, 1995; Richter *et al.*, 2000; Conard *et al.*, 2004; Conard *et al.*, 2009) and France (especially Isturitz) (Buisson, 1990; Scothern, 1992; Le Gonidec *et al.*, 1996; Lawson & d'Errico, 2002). These are clearly the product of complex production processes requiring considerable investment of time and resources, creating highly effective sound-producers, and must be the product of a long period of technological development (Dauvois, 1989, 1999; Lawson & d'Errico, 2002). Comparable pipes are found from contexts throughout the subsequent Upper Palaeolithic of Europe, as well as other artefacts that may have constituted sound-producers, including possible rasps and 'bullroarers' (Dauvois, 1989, 1999; Huyge, 1990, 1991). It also seems to be the case that Upper Palaeolithic *H. sapiens* made deliberate use of the acoustic properties of cave sites, including the resonant properties of stalagmites and stalactites in some places (Glory, 1964, 1965; Dams, 1984, 1985; Reznikoff & Dauvois, 1988; Dauvois, 1989, 1999; Reznikoff, 2008).

It would seem to be the case that amongst many, at least, of the Palaeolithic populations of Europe musical activities constituted an important part of their lives, being far from a trivial leisure activity adjunct to their subsistence concerns. This evidence confirms that the capacities for, and practice of, musical activities were well established in humans at this time; the development of the capacities for these behaviours clearly extends far further back than the last 40,000 years of our own species.

The evolution of the physiology and neurophysiology for musical behaviours: fossil evidence

In considering the earliest foundations of musical behaviours in the human lineage, one is necessarily investigating the origins of the production and processing of complex vocalisations and muscular movements. Without these capabilities, the musical behaviours that all humans undertake would be impossible.

The ability to produce and perceive varied sequences of tones, moderated for pitch, intensity and contour, is a fundamental component of musical behaviours. In contrast to the prevailing trend in Western music of the last few hundred years, instruments (anthropogenic sound-producers) are not fundamental to musical production; the human body has the potential to constitute an excellent instrument in its own right, both melodic and percussive. Instruments constitute an accessory to these existing human capacities; the origins of musical behaviour would not have relied upon the invention of instruments. The study of the origins of the capacities for musical behaviours must therefore examine the evolution of the biological features that are used in such activities.

The principal tonal sound-producing apparatus possessed by all primates is the vocal tract, and in humans, over the course of our evolution, this has become an instrument *par excellence*, with the potential to produce a great diversity of sounds, and to communicate information in a variety of ways. Indeed, it is this *biological* instrument, possessed by all of us, which constitutes the principal tonal sound-producer in the musical activities of many 'traditional' societies (e.g. Johnston, 1989; Nettl, 1992; Breen, 1994; McAllester, 1996; Locke, 1996).

Reconstructions of vocal anatomy have been carried out on both australopithecines (the bipedal but otherwise rather ape-like predecessors of our own genus, *Homo*, living from c. 4 million years ago until 1.5 million years ago or less), and the various species of *Homo* (which first appears around 2.5 million years ago). The australopithecines studied so far show characteristics of anatomy related to vocalisation that are little different from those of the African apes today (gorillas and chimpanzees) (Laitman & Heimbuch, 1982; MacLarnon & Hewitt, 1999; Alemseged *et al.*, 2006). Changes away from an ape-like resting position for the larynx are first evident in *H. ergaster*, which possesses the first indications of a lower resting laryngeal position and increased supralaryngeal soundspace (Laitman & Heimbuch, 1982; Arensberg *et al.*, 1990), which are amongst several changes which can increase the range of sounds that can be produced, and control over them (Fitch, 2009; Arensberg *et al.*, 1990; Clegg, 2012), and which were probably initially instigated by a shift to fully upright human-like bipedal posture (Aiello, 1996; Spoor & Zonneveld, 1998). While this has the potential to allow the production of a larger range of vocal sound frequencies than the ancestral (and australopithecine) form, the true range of sound-producing capabilities of the *H. ergaster* anatomy is difficult to model. It has been suggested that this development was coupled with an increase in neurological control of airflow over the larynx, as indicated by the dimensions of the central nerve canal in the cervical vertebrae, permitting some increased control of the pitch, intensity and contour of sounds produced by the larynx (Frayer & Nicolay, 2000), though the relevance of this anatomy for vocal control has been contested (Fitch, 2009). *H. ergaster* appears not to have undergone any increase in control over the *duration* of exhalation relative to the ape-like condition, as indicated by thoracic vertebral nerve canal dimensions, so although able to produce a greater variety of sounds, it would have been limited in the control of the length of the utterances it could produce, as are other higher primates today (MacLarnon & Hewitt, 1999; see Morley, 2012 for a discussion of the differing positions of MacLarnon & Hewitt, 1999, and Frayer & Nicolay, 2000).

By the time of the last common ancestor of Neanderthals and modern humans, probably around 5-600,000 years ago, human-like thoracic innervation had emerged, allowing control over utterances of extended duration (MacLarnon & Hewitt, 1999), alongside a modern-human-like hyoid anatomy and position, and thus supralaryngeal soundspace. Certainly European *H. heidelbergensis* specimens ancestral to Neanderthals, and African *H. heidelbergensis* specimens ancestral to modern humans, as well as Neanderthals and modern humans themselves, all possessed all of these features (Arensberg *et al.*, 1990; Rodríguez *et al.*, 2003; Martinez *et al.*, 2008). So a re-arrangement of laryngeal anatomy into a form essentially indistinguishable from that of modern humans, along with the neurological control over pitch, intensity, contour and duration of sounds produced by it, appears to have taken place at some point(s) over the 1-million-year or so period of the evolutionary development of *H. erectus*, from *H. ergaster* to the common ancestor of Neanderthals and ourselves.

That an increase in control over pitch, intensity and contour seems to have occurred before the ability to produce vocal sounds of extended duration is interesting. As MacLarnon & Hewitt (1999) point out, many primates vocalise in the form of discrete units of sound created with single air movements, but are limited in the duration of these and the order in which certain sounds can be made in the breathing cycle. They are also limited in the diversity of such sounds that they can make. An evolutionary path in which the ability to produce long sequences of controlled vocalisations developed out of an initial ability to make discrete vocalisations which were controlled for pitch and tone would seem to be consistent with the foundations for these capabilities which are already evident in higher primates.

On the basis of the available evidence, it seems likely that increasing control of intensity, pitch and intonation patterns of discrete vocalisations occurred initially, to date first exhibited by *H. ergaster*; pitch and intonation control increased subsequently with the continued development of a greater supralaryngeal soundspace, and control over maintaining long sequences of such utterances also followed, until these levels of control over vocal range and duration were essentially modern-like in *H. heidelbergensis*. It is possible that the ability to control extended sequences increased at the same time as vocal range increased, but the resolution of the record does not, at present, allow us to identify intermediate phases of either development – only where they both start (with *H. ergaster* at least c. 1.7 million years ago (m.y.a.), or an as yet undiscovered predecessor) and where they both appear complete (with *H. heidelbergensis*-like hominins c. 600,000 years ago). In fact, this sequence of the emergence of control, as suggested by the fossil evidence, makes more sense than the reverse – it is difficult to imagine how long sequences of vocalisations with little control over pitch, contour and intensity could be as meaningful as short sequences of vocalisations controlled for pitch, contour and intensity. The latter could be communicative in their own right, and as control increased, the length of sequences of such pitched and contoured utterances could also increase; subsequently, the order in which the expressive vocalisations occurred could assume importance.

The major changes in the vocal apparatus that can be tracked with reasonable confidence in the fossil record nevertheless have to be understood in terms of other changes in functionally-related neurological systems and behavioural capabilities in great apes and humans. Some insights into these processes can be gained both from comparative studies of contemporary neurological structures, and their relationships, in humans and primates, and fossil evidence for brain evolution. Fossil endocasts of hominin brains show particular development of regions in the left hemisphere, around Broca's area (Tobias, 1987; Bruner & Holloway, 2010), that are associated with fine muscular control of sequences of vocalisation and manual muscular movements (Ojemann *et al.*, 1989; Calvin, 1996; Duffau *et al.*, 2003; Nishitani *et al.*, 2005; Petrides *et al.*, 2005; Sergent *et al.*, 1992; Platel *et al.*, 1997; Besson & Schön, 2003; Mohr *et al.*, 1978; Poeppel & Hickock, 2004; Cantalupo & Hopkins, 2001). The earliest notable development of this area relative to australopithecines occurs with *H. habilis* and *H. rudolfensis* (Tobias, 1987; Bruner & Holloway, 2010), and the development of endocranial width at this point continues in subsequent hominins, being especially pronounced (non-allometrically) in Neanderthals and *H. sapiens* (though the extent to which this is disproportionate (non-allometric) in *H. ergaster* and *erectus* is equivocal) (Bruner & Holloway, 2010). It is important to note also that such changes in morphology and relative proportions of brain structures can be the consequence not only of changes in neurological structure but also of the mechanical and developmental constraints that exist upon cranial form which can in turn constrain the shape of the brain within (Bruner, 2004; Neubauer *et al.,* 2009; Bruner & Holloway, 2010). Nevertheless it is evident that in the case of *H. sapiens* the parietal regions in particular (and perhaps only these) have seen conspicuous non-allometric development; amongst other things these regions are involved in social communication, multi-modal processing, and the manipulation and planning of complex motor sequences (Bruner, 2004), all of which are critical elements of musical activity.

In primates, including humans, the motor planning of extended, purposeful utterances relies on input to the motor cortex from the ventral premotor and prefrontal cortex, including Broca's area (Jürgens, 2002; Cantalupo & Hopkins, 2001; Petrides *et al.*, 2005), and the *voluntary* integration of emotional content into vocalisations relies on input from the anterior limbic cortex and the periaqueductal grey matter (PAG) (Jürgens & Zwirner, 1986; Jürgens, 1992; Davis *et al.*, 1996; Schulz *et al.*, 2005). The PAG is also involved in reinforcing positive emotional experiences, including attachments to conspecifics and their vocal characteristics (Panksepp, 1995; Panksepp & Trevarthen, 2009). The nucleus ambiguus, which is directly adjacent to Broca's area, is responsible for integrating vocal fold control, expiratory control, orofacial muscular control and overall control of the laryngeal system (Vanderhorst *et al.*, 2001). But of the higher primates alive today, only humans possess the neurological connection allowing us to regulate the sound produced by the larynx itself, in combination with the use of our orofacial articulators and respiratory control (Jürgens, 1992; Jürgens, 2002; Schulz *et al.*, 2005; Okanoya & Merker, 2007). In doing this, however, we still rely on input from the mechanisms that organise reflex-like vocalisations in other primates

(Schulz *et al.*, 2005), reinforcing the idea that human vocal behaviour, although unique today amongst higher primates in its degree of voluntary control, built upon the existing system for vocalisations communicating emotional state and arousal.

So the ability to perform emotional vocal expression involving orofacial control and laryngeal activation, in response to external stimuli and internal affective state, seems to have been present in all primates on the lineage between rhesus monkeys and humans, but unlike other primates we are capable of vocal behaviour which involves voluntary control of the larynx, voluntary control and planning of the *structure* and *complexity* of vocal utterances, and a capacity for learning complex vocal patterns by imitation *and by invention* . Over the course of our evolution we have developed the monosynaptic neurological pathways necessary for this control, most likely since our divergence from the essentially ape-like australopithecines, and before our last common ancestor with Neanderthals, around 600,000 years ago.

Some features of auditory perception are obviously very ancient, being present in mammalian audition generally. One of these is the preferential perception of the so-called 'natural auditory categories'; these are also universal features of human speech sounds (Kuhl, 1988). This suggests that these qualities of vocalisation were tailored to the capabilities of the auditory system: as hominins developed the ability to control their vocalisations in order to communicate, there would have been strong selective pressure to be able to vocalise using these sounds that are most easily perceived by others. It would seem that audition, specifically the existence of 'natural auditory categories', was initially responsible for the formation of particular vocalisation properties, and that the mechanisms for perceiving such phonemic categories were in place in our hominin ancestors long before they were capable of actually producing articulate linguistic speech.

By contrast, other features of human auditory function appear to have faced significant selective pressure as a consequence of hominin vocalisation capabilities. For example, the human primary auditory cortex produces the greatest electrophysiological response to sounds in the 400Hz-4KHz frequency range, which is the range most useful for perceiving human speech sounds (Liégois-Chauvel *et al.*, 2003). At the physiological level, in marked congruence with the physiological evidence discussed above regarding the evolution of vocalisation anatomy, the anatomy of the middle and inner ear of hominins first shows significant changes towards a human-like form with *H. ergaster* (Spoor *et al.*, 1994; Spoor & Zonneveld, 1998). As with the vocalisation anatomy, this inner ear anatomy seems to be essentially modern-like by the time of *Homo heidelbergensis*: Martinez *et al.* (2004) show that this species (on the basis of five specimens from Sima de los Huesos, Atapuerca, Spain) also possessed middle ear anatomy which, like that of modern humans, was especially sensitive to the range of sound frequencies that are particularly salient in human speech vocalisations. Furthermore, the stapedius muscles of the middle ear in humans contract to reduce the movement of the stapes bones on the eardrum during vocalisation, and thus reduce the intensity of perception of our own vocalisations. This reduces the extent to which our own vocalisations obscure prevailing environmental sounds (Borg & Counter, 1989), and this ability would have become increasingly important as the length and range of vocalisations increased with the evolution of vocal anatomy. It is clear that aspects of our voluntary vocal sound-production capabilities and our auditory perceptual capabilities faced important selective pressures to co-evolve with each other in the context of maximising information extraction from these stimuli, and that these were essentially modern-like by the time of *H. heidelbergensis*, and thus likely our last common ancestor with Neanderthals.

The above and other evidence (for more detail see Morley, 2012, 2013) indicates that the possession of a vocal tract anatomy capable of producing sounds of variable pitch and extended duration has a very ancient evolutionary heritage. Rationales for the evolution of the human vocal tract have to account for the fact that it has developed in such a way that it allows us not only to produce a greater range of sound frequencies, but also to have very fine control over the entire range of those frequencies. They also have to account for why we are so sensitive to these frequency variations in utterances (their prosodic content). These elements of our vocalisation capabilities have very important communicative roles.

A fuller understanding of the emergence of vocalisation capabilities requires that we look at other sources of evidence too, and establish what their mutual implications are. What other evidence is there for control over the vocal system? And what is the use of vocalisations which are controlled for pitch and contour?

Tonal communication and speech: relationships and evolution of the neurology for the production and perception of vocal communication

Whilst humans have developed the specialised ability to voluntarily control the duration, structure and complexity of vocalisations, with precise control of the larynx and orofacial musculature, the process of vocalisation nevertheless relies on activation of deep-rooted and evolutionarily ancient instinctive emotional motor control neurology used in all primate vocalisations. These systems are involved in human vocalisations of all types (Jürgens, 2002; Schulz *et al.*, 2005; Snow, 2000).

In humans today the *production* of both vocal melodies and speech contours expressing emotion and intention (speech prosody) draw upon related structures (concerned with affective-tonal vocal production) (see above), but semantic elements of linguistic speech draw upon different, specialised, neurological structures (related not to the physical act of carrying out the vocalisation, but to the expression and comprehension of its meaning) (Marin & Perry, 1999).

Similarly, the *processing* of tonal content in both speech and music seems to rely on the same structures as each other: it appears that the ability to discriminate intonation patterns in speech (prosody) uses the same pitch discrimination mechanism as is used for pitch processing in music (Zatorre *et al.*, 1992; Patel *et al.*, 1998; Brust, 2003), but that the use of this mechanism by music is very refined, more refined than modern linguistic speech requires (Ayotte *et al.*, 2002). These mechanisms are located predominantly in the right hemisphere temporo-parietal region, in the superior temporal gyrus and frontal areas, with neurons in the right auditory cortex being especially tuned to pitch perception (Zatorre, 2003). The analysis of emotional tone content in speech seems to rely on activation in the right inferior frontal lobe, as well as evolutionarily ancient sub-cortical structures in the right hemisphere which are also used for processing emotional content in facial expression (Karow *et al.*, 2001; Belin *et al.*, 2004).

In terms of processing, whilst neurons in the right auditory cortex seem to be especially sensitive to spectral (tonal) information in auditory stimuli, those in the left auditory cortex seem to be especially sensitive to temporal information (Zatorre, 2003). Left hemisphere areas are also implicated in the capacity to perform planned sequences of complex muscular movements of rhythmic behaviour. These are important functions of Broca's area and the areas around it in the left hemisphere (see above), and these functions also form an important component of oral/praxic ability (Alcock *et al.*, 2000). The left hemisphere appears to be dominant with regard to semantic verbal meaning and syntactic sequencing and relationships (Benson, 1985); phoneme analysis relies on activation in the left inferior frontal lobe (Buchanan *et al.* 2000), and current anatomical evidence suggests that linguistic processing relies also on some input from sub-cortical structures in the left hemisphere (Karow *et al.*, 2001; see also Schulkin, 2013).

So both music and language functions use both left- and right-hemisphere structures; certain sub-functions of music and language seem to be shared, whereas functional lateralisation does seem to be the case for others (e.g. Borchgrevinck, 1982; Schweiger, 1985; Marin

& Perry, 1999; Brust, 2003). In particular, areas in the right hemisphere appear to be responsible for processing and production, in both melody and speech vocalisation, of prosodic melody, pitch control, tonality of singing, timbre processing and voice recognition (e.g. Benson, 1985; Bogen, 1985; Brown *et al.*, 2006; Brust, 2003). Left hemisphere regions appear to be implicated in production and processing of semantic verbal meaning and syntactic sequences, as well as rhythmic production and perception, planning and executing complex muscular sequences, and some aspects of conscious auditory analysis (e.g. Benson, 1985; Falk, 2000; Karow *et al.*, 2001; Besson & Schön, 2003; Brown *et al.*, 2006).

It is important to note that whilst some of the structures involved in specific aspects of auditory processing appear to be specifically lateralised to the left or right hemisphere, the overall process of sound perception involves activation of structures in both hemispheres, and in some cases specific tasks themselves also involve bilateral activation, albeit with some degree of bias towards greater activation in one hemisphere or the other. As Trevarthen (in press) emphasises, the specialisations exhibited by the two hemispheres (as evidenced by scanning technologies and studies of the effects of commissurotomy, where the connections between the two hemispheres in the corpus collosum are severed) have to be understood in terms of how they in fact work together. “In the normal brain of an individual the hemispheres work in tight partnership. The intuitive response to experience and the evocation of imagery linked to phenomenal reality by metaphor, on the one hand [right hemisphere functions], and verbal analysis and prescription of aesthetic judgements, on the other [left hemisphere functions], are two natural brain systems that develop as complements in the making of language, technology and art” (p. 10). It is in the contexts of their working together that cultural experience can be mastered, allowing effective participation in social life (Trevarthen, in press). This is particularly manifest in the case of musical activities: “…the lexicon of speech is limited for representation of the quality of imagination, and it seeks aid from expressive gesture, intonation and metaphor. Musical communication, which is universal among humans, serves to express affective relations and to establish a sense of belonging to a community of vital agents who share emotional appraisals of companionship in experience, from infancy (Malloch & Trevarthen, 2009[b]).” (p. 11).

The process of instigating vocalisation draws upon deep-rooted structures involved in tonal-emotional expression, and the perception of emotional content in tonal information also involves structures that are used for extracting emotional information from other sensory signals – i.e. other modes of emotional expression (Karow *et al.*, 2001). The specialised functions involved in linguistic verbal meaning have emerged later than these systems, apparently building upon some of the same structures in the left hemisphere that are required for the performance of planned sequences of complex muscular movements, including both vocalisation (through laryngeal and orofacial muscular control) and rhythmic behaviours (Alcock *et al.*, 2000; Besson & Schön, 2003).

So it is far from clear that any of the neurological structures that are used in processing the various aspects of musical stimuli are uniquely dedicated to that purpose, although it certainly appears that at least some of these structures have become finely tuned to the considerable processing demands of musical stimuli (Marin & Parry, 1999; Zatorre, 2003; Brown *et al.*, 2006). Vocal tonal production and the processing of tonal information each use a combination of both evolutionarily ancient structures involved in primate emotional vocal signalling, and structures which (whilst also used for other forms of communication) have become finely-tuned to the demands of musical activity. The structures that are used for music production and processing are also used in producing and processing aspects of other forms of communication, but this combination of neurological structures, and the interaction between them in producing and processing musical signals, represent a perhaps uniquely specialised combination of use of those mechanisms in the context of musical activities.

Indeed, following their review of a large body of research, Marin & Perry (1999) proposed that "The close correspondence between the networks of regions involved in singing and [linguistic] speaking suggests that [linguistic] speech may have evolved from *an already-complex system for the voluntary control of vocalisation*. Their divergences suggest that the later evolving aspects of these two uniquely human abilities are essentially hemispheric specialisations" (1999, p. 692 – emphasis added). The extent to which music and language processing overlap and share neural resources in adults and children (Patel, 2003; Koelsch *et al.*, 2003; Schön *et al.*, 2004; Koelsch *et al.*, 2005) lead Koelsch and Siebel to conclude that "it appears that the human brain, at least at an early age, does not treat language and music as strictly separate domains, but rather treats language as a special case of music." (2005, p. 582).

The fact that the various elements of musical activities draw upon cognitive mechanisms that are also used in similar ways during other activities – or *vice versa* – does not undermine the importance of musical activity in an evolutionary perspective, its relevance to the development of human cognition, or its importance in human behaviour; on the contrary these overlaps can emphasise its fundamentally important relationship with other critical aspects of human cognition and behaviour. This in itself has great implications for the role of evolution in the shaping of musical capabilities and the role of musical capabilities in the evolution of other aspects of human behaviour. We can look more closely at some of the overlaps between aspects of musical processing and the processing of other sound information, including speech; these relationships may provide some insight into how and why the functions emerged and developed.

Innate capabilities and non-linguistic vocal communication: some insights from developmental psychology and primate vocalisation

Human infants are born with abilities fundamental to musical processing, including the perception of frequency, timing and timbre, and are able to extract different emotional content from vocalisations, on the basis of tone and rhythm alone (Fernald, 1989, 1992; Trehub, 2003). In the emotional state that they express, some vocal sounds and frequency changes are fundamental, invariant across cultures, and even species (Morton, 1977, 1994; Scherer, 1985, 1986; Trainor *et al.*, 2000; Greiser & Kuhl, 1988; Fernald, 1992b, 1993; Werker *et al.*, 1994; Kitamura *et al.*, 2002). Part of the reason for this is that facial expression has a fundamental influence on vocal quality, as orofacial musculature helps determine properties of vocalisation such as frequency and vowel duration (e.g. Tartter, 1980; Falk, 2004b). This correlation between facial expression and vocal quality is shared by our nearest primate relatives, and similar correlations between vocal sound and emotional expression are also exhibited by several other species (Morton, 1977; Falk, 2004b; Bermejo & Omedes, 1999); this association evidently has an evolutionarily ancient provenance. Given the universality and innateness of certain fundamental facial expressions and the correspondence between these and characteristics of vocalisations, we can also expect characteristics of particular emotional vocalisations to be universal and innate too. We use facial affect and vocal affect to inform about the content of each other, inter-dependently within both production and perception (DeGelder & Vroomen, 2000; Belin *et al.*, 2004).

The universality of the vocal sounds and frequency changes that express particular emotions is especially evident in infant-directed (ID) speech, where the exaggeration of these elements of the vocalisation is a characteristic feature. ID vocalisations can tell us a great deal about the nature and role of the prosodic elements of speech and their relationship to musical melodic behaviour, as many of the properties of ID speech are shared with music. There are numerous parallels in terms of variable pitch contour, high rhythmicity, repetitive motifs, and the communication of affect, modulation of arousal, and eliciting of attention and affective response (e.g. Fernald, 1992a,b, 1993; Trehub *et al.*, 1993; Werker *et al.* 1994; Papousek, 1996; Lewkowicz, 1998;

Dissanayake, 2000; Mang, 2000; Trainor *et al.*, 2000; Falk, 2004b). It should be noted that the characteristic features of ID speech are also characteristics of the tonal (non-linguistic) elements of adult-directed (AD) speech, and they apparently share the same foundations and roles in emotional expression (Trainor *et al.*, 2000; Lavy, 2001). Vocalisations produced by pre-school children themselves are often difficult to classify as either linguistic or musical (Mang, 2000).

It seems that the best explanation for the shared prosodic pitch and tempo-related properties of emotional speech (ID and AD) and music is not that they are borrowed from one to the other, in either direction, but are, and always have been, a shared fundamental component of both. The music-like characteristics of ID vocalisations act upon cognitive-perceptual mechanisms that respond emotionally to emotional cues, and the characteristics of musical stimuli act upon the same mechanisms, not because these responses of the perceptual mechanisms are perpetuated into adulthood, but because their *function* is to respond in this way to these cues in *all* vocalisations, ID and AD. This emotional response to emotional cues is the foundation of empathy and successful interpersonal interaction, and musical stimuli act upon the mechanisms responsible as a reified form of the cues inherent in human emotional interaction. Response to such cues is as essential in adulthood as in childhood and infancy – but the use of those cues towards infants is more pronounced as a consequence of the need to develop and nurture those all-important interactive skills. Both music and planned use of vocalisation (especially ID but also AD) make use of an innate set of emotional responses to particular properties of vocalisations – properties of vocalisations that were extended from the communicative vocalisation activities of our primate and later hominin ancestors.

Already important in forming and maintaining social relationships amongst other primates (e.g. Elowson *et al.*, 1998a,b; Richman, 2000; Geissman, 2000; Gros-Louis, 2002), and likely our last common ancestor with them, it has been argued that such vocalisations increased in importance in our lineage initially due to the need to increase the efficiency of physical grooming activities (e.g. Aiello & Dunbar, 1993; Dunbar, 1998; Kudo & Dunbar, 2001), subsequently coming to be used in broader spheres of behaviour, though nevertheless maintaining a core role in the formation and maintenance of social relationships – as, indeed, full linguistic speech still does today (Emler, 1992; Dunbar *et al.*, 1997; Dunbar, 1998). In the case of four primate genera (*Indri*, *Tarsius*, *Callicebus* and *Hylobates*) 'singing' vocalisations additionally have a role specifically in reinforcing existing pair-bonded relationships, with both males and females singing and sometimes duetting (Geissman, 2000). Note that these behaviours do not seem to be used to 'woo' prospective mates, but occur between the members of an established mating pair. It seems that such activities are correlated with activities that *increase* pair bonding; Geissman & Orgeldinger (1998) observed that in ten siamang groups 'duetting activity was positively correlated with grooming activity and behavioural synchronisation, and negatively correlated with interindividual distance between mates' (Geissman, 2000, p. 111), suggesting that the activity is indeed related to the strength of pair bonds. Interestingly all known singing primate species are thought to have a monogamous structure, and this also applies to those bird species that duet (Geissman, 2000). To Geissman, 'This suggests that the evolution of singing behaviour in primates and of duet singing in general are somehow related to the evolution of monogamy' (p. 112); i.e. to the maintenance of a strong pair bond with a single other individual of the opposite sex – although its positive pair-bonding effects need not be limited to this context. Because the four groups of primates that sing are not closely related, in Geissman's view it is likely that this trait evolved independently four times in these genera (through convergent evolution), and he suggests that the same happened in hominins subsequently.

Several authors have proposed ways in which linguistic communicative systems may have emerged from such earlier forms of vocal communication. Linguistic speech and melody share common features in the form of intonational phrasing and combinatorial syntax, and share a 'phonological level' of meaning, an 'acoustic mode' involving emotive meaning and interpretation (Brown, 2000) – what might be termed 'intonational semantics'. Whilst music often does use the full range of sounds producible by the vocal tract, full language does not need to use the whole range to communicate effectively, as linguistic structures provide an additional source of semantic content and disambiguate meaning. This would not have been the case for pre-syntactic, pre-lexical forms of vocal communication. Tonal-contoured units expressive of affective state could partition into more discrete, smaller units (Wray, 1998) or become combined into progressively larger globally-contoured units (Brown, 2000), or both. In either case, what must have emerged in the course of the evolution of pitched-contoured vocalisations is an *increased range and increased control* of pitch contour, allowing greater vocal versatility, expressiveness, and thus efficiency, in proto-linguistic vocal affective communication. This need not have emerged as a new system of vocalisation initially, however; instead, it probably built upon the type of limited pitch control already used for emotive-tonal-social vocalisation amongst higher primates. The selective advantages associated with the possession of such capabilities, such as the formation of optimal co-operative, mating and parent-infant relationships would have resulted in the continued refinement of such capabilities through the evolution of the vocal tract and control over it, through the lowering of the larynx and increased innervation of the associated laryngeal and upper vocal tract musculature described above.

In these circumstances, the socially-important emotionally-communicative elements of such vocalisation would have remained the dominant element initially, with iconic and then abstract (symbolic) lexical associations subsequently increasing in importance. It is the latter – iconic and abstract lexical content – which relies on symbolic and analogical capacity and probably would have been a late-emerging element of communication, perhaps not prior to *H. sapiens*; the former – emotionally-communicative elements – do not rely on symbolic capacity, and would have been beneficial at a much earlier time. Even with the emergence of full lexical and syntactic language, the social-affective communicative foundations of prosodic contour remain a fundamental element of vocal communication – the same elements of vocalisation that form a foundation of melodic musical behaviours.

Gesture, timing and emotional communication: the integration of bodily gesture, vocal gesture, rhythm and emotion

Facial expression, or facial gesture, and vocalisation, or vocal gesture, are two facets of an interrelated system for the physical expression of emotion. This system also incorporates corporeal gesture, or 'body language'. Gesture and speech are inter-dependent. Both are affected simultaneously in stutterers (Mayberry & Jaques, 2000), and gestural and vocal behaviours are interrelated from the earliest babbling in infants (Locke *et al.*, 1995; Messinger & Fogel, 1998; Trevarthen, 1999; Butcher & Goldin-Meadow, 2000; Masataka, 2000; Falk, 2004b). They can operate independently, but when the upper limbs are otherwise unoccupied they are sequestered into speech-related gesture (Mayberry & Jaques, 2000). Furthermore, it seems the perception, as well as production, of vocalisation can be linked with gesture, and this is true from birth (Trevarthen, 1999).

Common timing processes are involved not only in movements of the limbs, but also in speech and non-speech movements of oral structures, suggesting that there is a cognitive rhythmic motor coordinator that instigates such muscular sequences irrespective of the musculature that is used, and that the complex patterns of muscular gesture (in fingers, hands, arms, shoulders and joints) and in vocalisation (orofacial, laryngeal and respiratory musculature) are co-ordinated (Franz *et al.*, 1992; Mayberry & Jaques, 2000). The concordance between gesture and speech is instigated early in the vocalisation process, by cyclical motor control, with gestural movements being associated with the nuclei of tone groups – prosodic rhythm – rather than the lexical elements of speech (McClave, 1994; Nobe, 1996; Mayberry & Jaques, 2000). This is also evident in the gestures accompanying infant vocalisations. These earliest gestures are emotive and rhythmic rather than iconic (which accords with the finding that gesture corresponds with prosodic rhythm rather than lexical content) (Trevarthen, 1999; Falk, 2004b) and only start to be used iconically (and in combination with words) when lexical behaviour has started to develop (Messinger & Fogel, 1998; Butcher & Goldin-Meadow, 2000).

Production of complex vocalisation relies on priming of the whole motor system (Blount, 1994; Feyereisen, 1997). Particular non-lexical vocal utterances (and non-linguistic content of speech) are accompanied by quite specific involuntary body-language; they share an affective origin in a system in which vocalisation and corporeal expression, or to put it another way, *vocal and corporeal gesture*, are intimately linked (Malloch & Trevarthen, 2009a). Note that this is at least equally applicable to body movements accompanying musical vocalisation, and vocalisation accompanying dance. There are not only shared neurological foundations between the ability to execute vocal and manual gestural sequences, but also a link between vocal and manual *rhythmic* capability (Franz *et al.*, 1992; Mayberry & Jaques, 2000). The capacity to perform rhythms, both manually and verbally, forms an important component of oral/praxic ability – detriments to one result in detriment to the other (Alcock *et al.*, 2000) – and this integration occurs whether utterances are linguistic or not (Franz *et al.*, 2000).

Although the production and perception of rhythm and melody involves some neurologically specialised and distinct areas of the brain which are in some respects independent of each other, there is also clearly important integration of these systems, with rhythmic muscular movements being coordinated with prosodic elements of vocalisation in their production (McClave, 1994; Nobe, 1996; Mayberry & Jaques, 2000). Specifically, they are inter-dependent in the planning and execution of sequences of muscular movement associated with instigation of vocalisation, rhythmic physiological movement, and expression of emotional state in these media. The production and perception of tonal *content* in vocalisations do not appear to require any input from rhythm-controlling systems in the left hemisphere, but the planning and execution of the muscular sequences themselves do.

Affective (emotional) content can apparently be interpreted equivalently in visual, auditory and kinaesthetic media, each of which can represent tension, release and particular emotions, underlining the cross-modality of affective expression and interpretation (Krummhansl & Schenck, 1997; Trevarthen, 1999; Mitchell & Gallagher, 2001). As noted, vocal quality is directly influenced by facial expression (Tartter, 1980; Tartter & Braun, 1994; Falk, 2004a,b), and the production of particular facial expressions and particular body postures actually causes us to experience some emotional response as a consequence (Levenson *et al.*, 1990). Furthermore, we tend to do this to some extent automatically when *witnessing* facial expressions and body language in others (Carlson, 1994; Wild *et al.*, 2001), and whether or not we fully physically manifest those expressions or postures ourselves, it is likely that action-observation-network neural firing occurs which replicates some of the brain response to adopting such a posture ourselves.

There seems to be a close interrelationship between the emotional-controlling elements of the limbic system and the areas responsible for the co-ordination of motor sequences and posture – the same systems that select and control movements also cause changes in the emotion-controlling elements of the limbic system (Trevarthen, 1999; Panksepp & Trevarthen, 2009). This can result in a self-directed feedback from movement into emotional state and, importantly, feedback and interaction between individuals, in terms of synchrony of movement and of emotional state. In other words, this physiological-emotional feedback may occur not only during production, but also during *perception* of such a stimulus. This means that in producing, and even to an extent in perceiving, a particular sound we generate some emotional response in ourselves due to the kinaesthetic feedback from the physiology and neurology required to produce that sound. It also means that we should expect there to be some consistent correlation across all humans between our emotional response to particular sounds, and the facial expression required to produce them.

The interrelationships discussed above mean that as well as being able to observe such cues, we can empathically experience something of others' emotional state in mirroring them with our own bodies. Musicality and rhythmic movement involve deliberate control and sequencing of this system, requiring us to adopt particular expressions and poses in the creation of the stimulus, to carry out particular vocal, facial and corporeal gestures, and furthermore they encourage the adoption of equivalent forms of these between individuals, which leads to a sharing of emotional state. Such reactions occur automatically whether one is fully participating in musical activity over all modalities, or only one (e.g. listening).

However important these inter-relationships are, solely producing or perceiving these stimuli is not enough to participate fully in musical activities. The ability to entrain – to synchronise movement with an internally- or externally-generated pulse – is a critical component of musical participation. Changes in frequency have direct effects on arousal level, and a consistent frequency can effectively moderate level of arousal too (Molinari *et al.*, 2003; Turner & Ioannides, 2009). Although they appear not to have the level of coordination to match beats exactly, infants engage in rhythmic movement in response to rhythmic stimuli, their movement rate correlating with rates in the stimuli, and their level of synchronisation positively correlates with positive emotional response (Zentner & Eorola, 2010; see also Provasi & Bobin-Begue, 2003; Kirschner & Tomasello, 2009). The ability to genuinely engage in entrainment may be directly related to the development of the abilities for sophisticated mimicry of gesture and corporeal expression, as well as having implications for abilities in turn-taking (critical in social interaction) and holding expectations about future events on the basis of patterns of events (Jones & Boltz, 1989; Bispham, 2006; Clayton *et al.*, 2005; Patel *et al.*, 2009).

The process of entraining may be *co-operative* or *subservient*, or some combination of the two (Clayton *et al.*, 2005), and has the potential to allow both 'losing oneself' in the stimulus, and/or a profound sense of physical cooperation, and synchronisation of arousal. Many musical experiences feature a powerful combination of both these effects, as individuals cooperatively (symmetrically) entrain their movements with each other whilst both entraining subserviently (asymmetrically) to music being played. The value of this experience is related also to the physicality of bodily gesture – the physical expression of emotional state – as well as direct overlaps in the mechanisms for the perception and production of musical stimuli, outlined above. Entrainment in the experience of music makes systematic use of our systems for understanding the emotional states and intentions of others through physical gesture, and is an inherent part of all musical experience.

It is tempting to intuitively view rhythm systems and melody systems as unrelated functions that have *come to be used together* in musical contexts over the course of the development of musical capabilities in humans, but this would seem to be inaccurate, artificially separating these systems. Whilst they clearly do rely on some specialised processing mechanisms, there are fundamental overlaps between them in that vocal control, rhythmic muscular movement, bodily gesture and emotional expression all rely on integrated systems which are activated in both production and perception of musical stimuli. Musical experience relies upon systematic use of a gestural system, including vocal tonal gesture relying on rhythmic cyclical muscular control, which exists to allow the expression of emotional state and the understanding of emotional state in others.

Music elicits emotional responses in listeners for a variety of reasons, in addition to the emotional reactions to tempo, prosody and gesture mentioned above. Musical stimuli can induce an emotional response both with and without conscious cognition of why it has done so, due to inherent properties of the music itself, and how it is processed (e.g. Juslin & Sloboda, 2001; Lavy, 2001; Steinbeis & Koelsch, 2008). Some emotional responses can be elicited as a consequence of learned associations with particular circumstances from our own experience, for example, others as a consequence of direct resemblance to ecological phenomena to which we have instinctive or conditioned responses (Juslin & Sloboda, 2001; Cross, 2003a). Properties of musical stimuli can elicit genuine physiological reactions equivalent to those elicited by emotional expression in other media, such as changes in respiration, heart-rate, temperature and tingling, and they are processed by many of the same mechanisms (McFarland & Kennison, 1989; Sloboda, 1991, 1998; Panksepp, 1995; Krummhansl, 1997; Nyklicek *et al.*, 1997; Panksepp & Bernatsky, 2002; Blood & Zatorre, 2001; Trainor & Schmidt, 2003). Some of these reactions are caused by neurochemicals that are related to the formation of social bonds, reduction of separation anxiety and seeking of reward and gratification (Blood & Zatorre, 2001; Trainor & Schmidt, 2003; Menon & Levitin, 2005; Panksepp & Trevarthen, 2009).

It will be clear from the preceding discussion that some of the very significant causes of emotional response to music are a consequence of the processing of music by, and its stimulation of, auditory and kinaesthetic mechanisms associated with interpersonal interaction. Musical stimuli can be interpreted as having human-like properties (Watt & Ash, 1998; Davies, 2001; Lavy, 2001), and can have similar effects to interacting with a person, through being processed using mechanisms related to the interpretation of meaning in interpersonal interaction. The contours of musical stimuli can have much in common with physical (including vocal) expression of emotional state, stimulating the interpretation of emotion across the other media that would normally be associated physically with that contour (Clynes, 1977; Scherer, 1991; Davies, 2001). The dynamic character of public physiological expression, and musical contour and tempo, are processed as part of the same system of expression with some of the auditory cues in music being interpreted *in the same way* as physiological and corresponding auditory expressions of emotional state (Wagner, 1989; Kappas *et al.*, 1991; Lavy, 2001).

We can react to such cues sympathetically, through recognising those emotional cues, or empathically, feeling a shared experience with the emotion detected, if it elicits the same emotion as is being expressed. There are strong associations between sociality, empathic ability, expressiveness and motivation to musicality, all being particularly prominent in Williams' syndrome individuals (Levitin & Belugi, 1997; Huron, 2001; Panksepp & Trevarthen, 2009), and often severely diminished in autistic individuals (e.g. Sloboda *et al.*, 1985; Davison & Neale, 1994; Huron, 2001; Sacks, 2008).

The context in which we experience music is also very important in determining the emotion, and intensity of emotion, experienced. Especially important in this respect is the social context – the extent to which the experience is shared with others, and their reaction to the same stimuli, with emotions being 'contagious' and self-reinforcing (Kraut & Johnston, 1979; Hatfield *et al.*, 1994; Wild *et al.*, 2001). Musical experience can gain much of its value from a sense of a profoundly personal response coupled with the sense of shared experience; meanwhile, when practised alone, it can act as a surrogate for interaction and shared experience.

So a wide range of evidence indicates that musical stimuli act upon, and derive from, evolved mechanisms in the human brain and body that are fundamentally related to each other, capacities for interpersonal interaction, imitation, learning, and the expression, comprehension and moderation of emotional state. These inter-dependent capacities, and the other, related, behaviours that they support, can be collectively termed *musicality* (see also Malloch & Trevarthen, 2009a). It would seem that it is more accurate to view musical behaviours as a specialised, dedicated context of use of this capacity for musicality, rather than as a post-hoc *appropriation* of these underlying abilities into a new, unrelated context.

Does musical practice itself confer selective advantages?

The preceding discussion has been very much concerned with the evolution and efficacy of these underlying capacities for musicality; we must also ask whether there are any reasons, selective or otherwise, why musical behaviours, as a dedicated context systematically exercising this capacity for musicality, may have become developed and perpetuated within hominin (and/or human) groups. In other words, are there, or were there at some point in the past, selective advantages to carrying out musical behaviours themselves?

It has often been remarked that music is hardly essential for survival, the implication being that there is no reason to believe that musical behaviours could have been selectively important. However, to confer a selective benefit a behaviour or trait need not be *essential* for survival, it need only confer a slightly greater likelihood of successful procreation, and thus greater rate of procreation (thus perpetuating that trait), than would otherwise be the case.

As has been noted, in talking about selection for musical behaviours we can either be talking about selection for carrying out musical activities, as a behavioural package, or we can be talking about selection for each of the foundations of musical abilities, which could have other selective pressures acting upon them as a factor of other functions that they fulfil. Past discussion of the potential values of musical behaviours with regard to selection have not always made this distinction between rationales for the evolution of musical capacities, and rationales for the perpetuation of developed musical behaviours; further, it is important to make a clear distinction between biological selection for musical behaviours, and non-biologically selective reasons why they may be perpetuated.

That musical and proto-musical behaviours use mechanisms that are selectively important in contexts other than their use in music does not diminish the importance of musical behaviours, from an evolutionary perspective, if music is not simply *making use of* existing cognitive mechanisms that already existed, selected for already, but is *a development of* those mechanisms, fulfilling the roles of those mechanisms in an additional context. Whilst those foundations may initially have been selectively favoured as a consequence of their fulfilment of particular purposes, music developed within the context of those uses, and musical behaviours have the potential to fulfil some of those same purposes – potentially in even more effective ways. Such a suite of related capacities could continue to develop in tandem, with interdependence increasing between them, whilst still fulfilling other functions. Such co-use of mechanisms previously related in their fulfilment of different functions could then unite them functionally in this new behavioural manifestation; subsequently they could be selected for in tandem as part of a behavioural system, changes in one mechanism ‘bootstrapping’ changes in others. As Huron (2001) puts it, “If music is an evolutionary adaptation, then it is likely to have a complex genesis. Any musical adaptation is likely to be built on several other adaptations that might be described as premusical or protomusical. Moreover, the nebulous rubric *music* may represent several adaptations, and these adaptations may involve complex co-evolutionary patterns with culture” (Huron, 2001, p. 44).

There are several ways in which behaviours, the capacities which support those behaviours, and evolution by selection may interact. In the case of musical behaviours, I suggest that these can be characterised as the following selective processes:

1) Selection for *capacities* underlying musical behaviours because of their value in other circumstances. Musical practice then gains its efficacy (effects and wider emotional and

social benefits) from its use of these capacities. This efficacy then may or may not itself have selective benefits, but this would be a separate process;

2) Selection for the *capacities* that support musical behaviours, through their use in music, because of benefits of *exercising those capacities together* for other aspects of life (i.e. the action of participating in musical activities itself *indirectly* facilitates individual survival and procreation);
3) Selection for the *capacities* that support musical behaviours in the context of their *efficacy in their use in music* – i.e. benefitting the practice of music itself, which is itself selectively advantageous for some reason (i.e. the action of participating in musical activities itself *directly* facilitates individual survival and procreation);
4) Sexual selection for musical *capacities* via musical *practices* because of the practices *indicating* fitness of participants due to potential survival benefits of the capacities that support them;
5) Cultural selection for particular musical *practices* (including capabilities to participate in those practices), which then may or may not have a biological effect through gene-culture co-evolution, through social functions of musical practice having knock-on effects to individual survival and reproductive rate.

Note that none of the five selective processes is mutually exclusive – any or all of them could potentially be acting during evolutionary history, at different times or simultaneously. The distinctions between these different selective processes have not generally been made in discussions of music and evolution; as a consequence some proposed models have conflated various aspects of them, some models have often been argued to be valid to the exclusion of others, and criticisms of them have often been on the basis of mutual exclusivity with other possibilities.

Process (1) includes rationales regarding the origins of musical behaviours in parent-infant interaction (e.g. Dissanayake, 2000), and in synchronous chorusing (Merker, 1999) for example. In addressing the question of selective roles for musical behaviours themselves, via processes (2), (3), (4) and (5), it remains to be asked whether there are genuine circumstances in which such mechanisms could have operated and, if so, how they would have influenced the development of musical behaviours.

It is possible that developed musical behaviours could provide a good medium for the use and display of various traits related to fitness, and that aspects of those behaviours might be 'fine-tuned' by sexual preferences exhibited under such circumstances (c.f. Miller, 2000). However, there are several shortcomings of such explanations offered so far. This mechanism cannot account for those preferences' existence, or for their being applicable to music to start with; these require other explanations. Further, a distinction needs to be made between cultural sexual selection and biological sexual selection: *behaviours* could conceivably be perpetuated as a consequence of sexual preference, but unless the behaviour is a consequence of a biological trait, which is then itself selected for, such perpetuation will be through social practice and not Darwinian sexual selection.

One way – with a biological basis – in which music could be a particularly potent expression of reproductive fitness lies in its roots in the ability to communicate emotionally and, effectively, to empathise, bond and elicit loyalty. An individual who is talented in these respects may well be more appealing to the opposite sex, because they are more likely to be able to form strong social alliances, and strong pair and family bonds. Good musical ability may vicariously indicate such abilities, as the cognitive capacities relied upon are in many respects shared.

The fact that people are frequently strongly drawn to same-gender music groups actually does not undermine the above appeals of music, as forming strong alliances is not a gender-specific activity; music can also fulfil the role of engendering strong feelings of empathic association and group membership (and thus loyalty and cooperativity are both more likely to be offered and reciprocated). Whilst musical behaviours may not directly indicate an individual's ability to contribute to a group in subsistence terms (Hagan & Bryant, 2003), such a criticism might equally be levelled at grooming activity, which is known to confer significant individual advantages. This fails to account for any benefit music might have in the respects discussed above, namely as an indication of an individual's *likelihood* of contributing to a group or to specific individuals, as indicated by their networks and loyalty, and as a mechanism for actually stimulating and maintaining those networks and loyalties.

Group musical activities may form an important signal (and stimulus) of group cohesion, cooperation and loyalty, and a forum for the coordination of complex interactions; it still frequently fulfils this role. A distinction again needs to be made between a "selective basis for the foundations of musical behaviours" and "reasons why such behaviours may have been perpetuated". Coalition signalling (Hagan & Bryant, 2003) is not a strong candidate for an explanation for the former, but it may well be a good explanation for the latter. Group displays of musical behaviours can indicate group stability and the ability to carry out complex coordinated actions precisely because they can *engender* these things.

Indeed, it is telling that even in the modern Western world, where our recent economic, social and technological changes mean that we are often separated from the producers of music, musical activities commonly remain firmly in the social sphere. This is even more evident in the case of contemporary and recent historical 'traditional' hunter-gatherer societies. For example, many of the most important roles of music amongst Plains Native Americans (Nettl, 1992; McAllester, 1996), African Pygmies (Kisliuk, 1991; Turino, 1992; Ichikawa, 1999; Lewis, 2009), Yupik and Inuit Eskimos (Nettiez, 1983; Johnston, 1989), and Australian Aborigines (Breen, 1994; Myers, 1999) are social, interactive, and integrative, and the participants themselves often see these as the most important consequences of the activity.

Musical behaviours could be valuable not only as a means of exploring social interactions, but actually as a vicarious stimulus and exercise of those capacities (see, for example, Cross, 2001). Musical or proto-musical behaviours also have the potential to make use of several domains of intelligence at once, relying on the integration and control of biological, psychological, social and physical systems; furthermore musical activities give the opportunity to practice and develop these integrated skills in a context of limited risk (Cross, 2001). Broader mechanisms of selection (Shennan, 2002; Bowles & Gintis, 1998; Boyd & Richerson, 2005) may prove fruitful in generating models of processes through which musical behaviours may have come to develop and spread, through the above advantages, conferred on individuals within groups as a consequence of the stability and effectiveness of their group.

Conclusions

It is tempting to attempt to identify a 'moment' when musical behaviours emerged as an activity in their own right, separable from other activities, but it is clear that music is not genuinely separable from certain other aspects of our lives, interactions and abilities. The elements of its production and perception, its forms and uses are integrally related to other aspects of our social existence. There was no single moment when music appeared, but we can hope to better understand how activities that we would recognise as musical emerged, and why they have the effects and roles that they do.

Vocal tonal expression and comprehension is a highly refined component of a system for gestural expression and comprehension of emotional state, which also involves body language, manual gesture and facial expression, and which relies on a cyclical rhythmic motor-planning mechanism for its execution. Musical activities, including dance, constitute a refined and 'distilled' version of the use of this system, with auditory modes of expression (including tone) being emphasised in music and corporeal modes being emphasised in dance.

This wider system always has encompassed, and continues to encompass, other communicative behaviours too, which share many overlapping features with music. These include the paralinguistic aspects of vocal communication, including prosody, and the vocalisations directed at pre-linguistic infants. Both these elements have clear foundations in the communicative behaviours of other higher primates, and over the course of the human lineage the extent of voluntary deliberate control over them, and with that their complexity, has increased.

The fossil record of human ancestors indicates that the vocal anatomy and neurological structure of australopithecines, the rather ape-like but bipedal predecessors of our genus, *Homo*, was not significantly different from that of the great apes of today. By the time of *H. heidelbergensis*, a descendent of *H. erectus* that probably also represents the last common ancestor of Neanderthals and *H. sapiens*, a vocal anatomy, and neurological control over it that looks essentially modern, appears to have been in place, along with modern-human-like auditory physiology, fine-tuned to human vocalisation frequencies. This species, by at least 500,000 years ago, would appear to have had the ability to finely control a range of vocal sound frequencies comparable to our own, and to do so in the context of utterances of extended duration. Whilst a range of factors are likely to have been responsible for the changes that made this possible, including the earlier shift to a fully upright bipedal posture, increasing brain size, and increasing demands on manual dexterity, it would seem likely that the combination of vocal expressive abilities made possible was also subject to positive selection in leading to this combination of features.

Amongst higher primates and humans today communicating social-emotional content remains one of the most important roles of vocal behaviour, and this is likely to have been the case for our hominin ancestors as well. Non-verbal vocal utterances, which are also coupled with equivalent body language and facial expression, can express personal state and reactions (wellbeing, approval, disapproval, disgust, etc.) and can also solicit such information from other individuals. Both the expression and solicitation of such information, as well as the ability to share in the experience of such states, are important and advantageous skills for forming and managing social relationships, including co-operative, mating and parent-infant relationships. Individuals most effectively able to establish and maintain pair-bonds and alliances through this ability would have a significant selective advantage over less able fellows.

The production and perception of gestural (vocal, orofacial and corporeal) expression of emotion in this system involves the priming of the rhythmic-motor and emotional systems. Rhythmic sequences, and the prosodic and rhythmic content of tonal sequences, prime this system and each other, resulting in a multimodal relationship between rhythmic and emotive tonal content. This takes the form of auditory, visual and kinaesthetic expression of emotion, and it would appear that what emerged over the course of the evolution of *Homo* was the ability to deliberately use this system, along with increasing control over the form, range and duration of these expressive gestures. It is proposed here that the culturally-shaped melodic, rhythmic behaviours that we call music, and semantic, lexical linguistic abilities, later emerged as specialised behaviours building upon the foundations of this system of vocal and kinaesthetic communication of emotion.

It is clear that by the time of the arrival of *H. sapiens* in Europe more than 40,000 years ago, these developments had taken their recognisable modern forms, and it is likely that the evidence of these behaviours that we find then does not constitute evidence of the first instance of these behaviours.

The preceding discussion, I would argue, justifies a definition of music and dance together, at their most fundamental level, as *deliberate metrically-organised gesture*. In this context *gesture* refers to the embodied expression of emotion (i.e. vocal and orofacial and/or corporeal motor action incorporating input from emotion-controlling systems), *metrically-organised* refers to the coordination of the gestures with an internally- or externally-generated temporal pattern, and *deliberate* refers to being under conscious control (note that undertaking musical activity involves many unconscious processes too, though the act itself is volitional). This definition is certainly reductionist, and there are many elements of the *experience* of music that it does not encompass, being individual- and culture-specific elements that are built around this core, but the fundamental, universal, essential elements of our musical experience are, I would argue, derived from the fact that what we are experiencing is deliberate metrically-organised gesture.

Whether or not musical activities themselves can or did confer selective advantages themselves, it is clear that the repertoire of capabilities constituting musicality, of which musical behaviours are a specialised part, are both rooted in ancient capabilities, developed and refined in humans through selection, and fundamental to many aspects of the uniquely complex manifestation of social structures that we see in humans today. It seems that these abilities are not only a ubiquitous feature of humanity today, but a fundamental component of what makes human societies possible.

Acknowledgements

I am grateful to Cristian Capelli for his invitation to submit this review, to Giovanni Destro Bisol and Emiliano Bruner for their editorial input, and to Emiliano and the two reviewers for their very helpful comments. The research was variously funded by The Wenner-Gren Foundation for Anthropological Research (New York), The Arts and Humanities Research Council (UK), and Trinity Hall, Cambridge.

References

Aiello L. 1996. Terrestriality, bipedalism and the origin of language. *Proc. Br. Acad.*, 88: 269-289.

Aiello L. & Dunbar R. 1993. Neocortex size, group size, and the evolution of language. *Curr. Anthropol.,* 34: 184-193.

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

Alemseged Z., Spoor F., Kimbel W., Bobe R., Geraads D., Reed D & Wynn J. 2006. A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature,* 443: 296-301.

Arensberg P., Schepartz L., Tillier A., Van DerMeersch B. & Rak Y. 1990. A reappraisal of the anatomical basis for speech in middle Palaeolithic hominids. *Am. J. Phys. Anthropol.,* 83: 137-146.

Ayotte J., Peretz I. & Hyde K. 2002. Congenital amusia. A group study of adults afflicted with a music-specific disorder. *Brain,* 125: 258-231.

Bannan N. 2012. *Music, language and human evolution*. Oxford University Press, Oxford.

Barrett L., Dunbar R. & Lycett J. 2001. *Human Evolutionary Psychology*. Palgrave Macmillan, Basingstoke.

Belin P., Fecteau S. & Bédard C. 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.,* 8: 129-135.

Benson D. 1985. Language in the left hemisphere. In D. Benson & E. Zaidel (eds): *The Dual Brain,* pp. 193-203. Guildford Press, New York.

Bermejo M. & Omedes A. 1999. Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungo (Democratic Republic of Congo). *Folia Primatol.,* 70: 328-357.

Besson M. & Schön D. 2003. Comparisons between language and music. In I. Peretz & R. Zatorre (eds): *The Cognitive Neuroscience of Music*, pp. 269-293. Oxford University Press, Oxford.

Bispham J. 2006. Rhythm in music: What is it? Who has it? And why? *Music Percept.,* 24: 125–134.

Blacking J. 1995. *Music, Culture and Experience: Selected Papers of John Blacking* (edited by P. Bohlman & B. Nettl). University of Chicago Press, Chicago.

Blood A. & Zatorre R. 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U.S.A.,* 98: 11818-11823.

Blount B. 1994. Comment on Armstrong, Stokoe & Wilcox. *Curr. Anthropol.,* 35: 358-359

Bogen J. 1985. The dual brain: some historical and methodological aspects. In D. Benson & E. Zaidel (eds): *The Dual Brain*, pp. 27-43. Guildford Press, New York.

Bohlman P. 2002. *World Music*. Oxford University Press, Oxford.

Borchgrevink H. 1982. Prosody and musical rhythm are controlled by the speech hemisphere. In M. Clynes (ed): *Music, Mind and Brain,* pp. 151-157. Plenum Press, New York.

Borg E. & Counter S. 1989. The middle-ear muscles. *Scientific American*, August 1989, 63-68.

Bowles S. & Gintis H. 1998. The moral economy of community: structured populations and the evolution of pro-social norms. *Evol. Hum. Behav.,* 19: 3-25.

Boyd R. & Richerson P. 2005. *The Origin and Evolution of Cultures*. Oxford University Press, Oxford.

Breen M. 1994. I have a dreamtime: aboriginal music and black rights in Australia. In S. Broughton, M. Ellingham, D. Muddyman & R. Trillo (eds): *World Music: The Rough Guides,* pp. 655-662. Rough Guides Limited, London.

Brown S. 2000. The "musilanguage" model of music evolution. In N. Wallin, B. Merker & S. Brown (eds): *The Origins of Music*, pp. 271-300. MIT Press, Cambridge, MA.

Brown S., Merker B. & Wallin N. 2000. An introduction to evolutionary musicology. In N. Wallin, B. Merker & S. Brown (eds): *The Origins of Music,* pp. 3-24. MIT Press, Cambridge, MA.

Brown S., Martinez M. & Parsons L. 2006. The neural basis of human dance. *Cereb. Cortex,* 16: 1157-1167.

Bruner E. 2004. Geometric morphometrics and paleoneurology: brain shape evolution in the genus *Homo*. *J. Hum. Evol.,* 47: 279-303.

Bruner E. & Holloway R. 2010. A bivariate approach to the widening of the frontal lobes in the genus *Homo*. *J. Hum. Evol.,* 58: 138-146.

Brust J. 2003. Music and the neurologist: a historical perspective. In I. Peretz & R. Zatorre (eds): *The cognitive neuroscience of music,* pp.181-191. Oxford University Press, Oxford.

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

Buisson D. 1990. Les flûtes paléolithiques d'Isturitz. *Bull. Soc. Préhist. Française,* 87: 420-433.

Burns E. 1999. Intervals, scales and timing. In D. Deutsch (ed): *The psychology of music* (2nd Edition), pp. 215-264. Academic Press, San Diego.

Butcher C. & Goldin-Meadow S. 2000. Gesture and the transition from one- to two-word speech: when hand and mouth come together. In D. McNeill (ed): *Language and gesture,* pp. 235–258. Cambridge University Press, Cambridge.

Butler D. 1989. Describing the perception of tonality in music: a critique of the tonal heirarchy theory and proposal for a theory of intervallic rivalry. *Music Percept.,* 6: 219-42.

Calvin W. 1996. *How Brains Think*. Weidenfeld & Nicholson, London.

Cantalupo C. & Hopkins W. 2001. Asymmetric Broca's area in great apes. *Nature,* 414: 505.

Carlson N. 1994. *The Physiology of Behavior*. Allyn & Bacon, Massachusetts.

Clayton M., Sager R. & Will U. 2005. In time with the music: the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology,* 11: 3–142.

Clegg M. 2012. The evolution of the human vocal tract: specialized for speech? In N. Bannan (ed): *Music, Language and Human Evolution, pp*. 58-80. Oxford University Press, Oxford.

Clynes M. 1982. *Music, Mind and Brain: The Neuropsychology of Music.* Plenum Publishing Corporation, New York.

Clynes M. 1977. *Sentics: the Touch of Emotions*. Souvenir Press, New York.

Conard N., Malina M. & Münzel S. 2009. New flutes document the earliest musical tradition in southwestern Germany. *Nature,* 460: 737-740.

Conard N., Malina M., Münzel S. & Seeberger F. 2004. Eine Mammutelfenbeinflöte aus dem Aurignacien des Geissenklösterle: Neue Belege für eine Musikalische Tradition im Frühen Jungpaläolithikum auf der Schwäbischen Alb. *Archäologisches Korrespondenzblatt,* 34: 447-462.

Cross I. 2003a. Music and Emotion: lecture notes, available online at *http://www.mus.cam.ac.uk/~ic108/IBMandS/emotion/musicemotion.html*

Cross I. 2003b. Music and biocultural evolution. In M. Clayton & T. Herbert (eds): *The cultural study of music: a critical introduction,* pp. 19-30. Routledge, London.

Cross I. 2001. Music, mind and evolution. *Psychol. Music,* 29: 95-102.

Cross I. & Morley I. 2009. The evolution of music: theories, definitions and the nature of the evidence. In S. Malloch & C. Trevarthen (eds): *Communicative Musicality,* pp. 61-81. Oxford University Press, Oxford.

Dams L. 1985. Palaeolithic Lithophones: Descriptions and Comparisons. *Oxford J. Archaeol.,* 4: 31-46.

Dams L. 1984. Preliminary Findings at the "Organ Sanctuary" in the Cave of Nerja, Malaga, Spain. *Oxford J. Archaeol.,* 3: 1-14.

Darwin C. 1871. *The Descent of Man and Selection In Relation To Sex.* John Murray, London.

Dauvois M. 1999. Mesures acoustiques at témoins sonores osseux Paléolithiques. *Préhistoire d'os, recuil d'études sur l'industrie osseuse préhistorique offert à Mme Henriette Camps-Febrer.* Publications de l'Université de Provence, Aix-en-Prevence.

Dauvois M. 1989. Son et Musique Paléolithiques. *Les Dossiers D'Archéologie,* 142: 2-11.

Davies S. 2001. Philosophical perspectives on music's expressiveness. In P. Juslin & J. Sloboda (eds): *Music and Emotion: Theory and Research,* pp. 23-44. Oxford University Press, Oxford.

Davis P., Zhang S., Winkworth A. & Bandler R. 1996. Neural control of vocalisation: respiratory and emotional influences. *J. Voice,* 10: 23-38.

Davison G. & Neale J. 1994. *Abnormal Psychology* (6th Edition). John Wiley & Sons, New York.

d'Errico F., Henshilwood C., Lawson G., Vanhaeren M., Tillier A.-M., 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. *J. World Prehist.,* 17: 1-70.

DeGelder B. & Vroomen J. 2000. The perception of emotions by ear and by eye. *Cogn. Emot.,* 14: 289-311.

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

Duffau H., Capelle L., Denvil D., Gatignol P., Sichez N., Lopes M., Sichez J.-P. & Van Effenterre R. 2003. The role of dominant premotor cortex in language: a study using intraoperative functional mapping in awake patients. *Neuroimage,* 20: 1903-1914.

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

DunbarR., Duncan N. & Marriott A. 1997. Human conversational behaviour. *Hum. Nat.,* 8: 231-246.

Elowson A., Snowdon C. & Lazaro-Perea C. 1998a. "Babbling" and social context in infant monkeys: parallels to human infants. *Trends Cogn. Sci.,* 2: 31-37.

Elowson A., Snowdon C. & Lazaro-Perea C. 1998b. Infant "babbling" in a non-human primate: complex vocal sequences with repeated call-types. *Behaviour,* 135: 643-664.

Emler N. 1992. The truth about gossip. *Social Psychology Newsletter,* 27: 23-37.

Falk D. 2004a. *Braindance: New Discoveries about Human Origins and Brain Evolution* (revised and expanded edition). University Press of Florida, Gainesville FL.

Falk D. 2004b. Prelinguistic evolution in early hominins: whence motherese? *Behavioural and Brain Sciences,* 27: 491–541.

Falk D. 2000. Hominid brain evolution and the origin of music. In N. Wallin, B. Merker & S. Brown (eds): *The Origins of Music,* pp. 197-216. MIT Press, Cambridge MA.

Fernald A. 1993. Approval and disapproval – infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Dev.,* 64: 657-674.

Fernald A. 1992a. Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens & M. Papousek (eds): *Nonverbal Vocal Communication,* pp. 262-282. Cambridge University Press, Cambridge.

Fernald A. 1992b. Human maternal vocalisations 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,* pp. 391-428. Oxford University Press, Oxford.

Fernald A. 1989. Intonation and communicative intent in mothers' speech to infants: is melody the message? *Child Dev.,* 60: 1497-1510.

Feyereisen P. 1997. The competition between gesture and speech production in dual-task paradigms. *J. Mem. Lang.,* 36: 13-33.

Fitch W. 2009. *Fossil cues to the evolution of speech.* In R. Botha & C. Knight (eds): *The Cradle of Language,* pp. 108-130. Oxford University Press, Oxford.

Franz E., Zelaznik H. & Smith A. 1992. Evidence of common timing processes in the control of manual, orofacial, and speech movements. *J. Mot. Behav.,* 24: 281-287.

Frayer D. & Nicolay, C. 2000. Fossil evidence for the origin of speech sounds. In N. Wallin, B. Merker & S. Brown (eds): *The Origins of Music,* pp. 217-234. MIT Press, Cambrige MA.

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

Geissman T. & Orgeldinger M. 1998. Duet or divorce! *Folia Primatol.,* 69: 283.

Glory A. 1965. Nouvelles découvertes de dessins rupestres sur le causse de Gramat. *Bull. Soc. Préhist. Française,* 62: 528-536.

Glory A. 1964. La Grotte de Roucador. *Bull. Soc. Préhist. Française,* 61: 166-169.

Gosselin N., Samson S., Adolphs R., Noulhiane M., Roy M., Hasboun D., Baulac M. & Peretz I. 2006. Emotional responses to unpleasant music correlates with damage to the parahippocampal cortex. *Brain,* 129: 2585-2592.

Greiser D. & Kuhl P. 1988. Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Dev. Psychol.,* 24: 14-20.

Gros-Louis J. 2002. Contexts and behavioural correlates of trill vocalisations in wild white-faced Capuchin monkeys (*Cebus capucinus*). *Am. J. Primatol.,* 57: 189-202.

Hagan E. & Bryant G. 2003. Music and dance as a coalitional signaling system. *Hum. Nature*, *14*: 21-51.

Hahn J. & Münzel S. 1995 Knochenflöten aus den Aurignacien des Geissenklösterle bei Blaubeuren, Alb-Donau-Kreis. *Fundberichte aus Baden-Würtemberg,* 20: 1-12.

Hatfield E., Cacioppo J. & Rapson R. 1994. *Emotional Contagion*. Cambridge University Press, Cambridge.

Higham T., Basell L., Jacobi R., Wood R, Bronk Ramsey C. & Conard N. 2012. Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: The radiocarbon chronology of Geißenklösterle. *J. Hum. Evol.,* 62: 664-676.

Huron D. 2001. Is music an evolutionary adaptation? *Ann. N. Y. Acad. Sci.,* 930: 43-61.

Huyge D. 1991. The "Venus" of Laussel in the light of ethnomusicology. *Archeologie in Vlaanderen,* 1: 11-18.

Huyge D. 1990. Mousterian skiffle? Note on a Middle Palaeolithic engraved bone from Schulen, Belgium. *Rock Art Research,* 7: 125-132.

Ichikawa M. 1999. The Mbuti of Northern Congo. In R. Lee & R. Daly (eds): *The Cambridge Encyclopedia of Hunters and Gatherers, pp.* 210-214. Cambridge University Press, Cambridge.

Johnston T. 1989 Song categories and musical style of the Yupik Eskimo. *Anthropos,* 84: 423-431.

Jones M. & Boltz M. 1989. Dynamic attending and responses to time. *Psychol. Rev.,* 96: 459-491.

Jürgens U. 2002. Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.,* 26: 235-258.

Jürgens U. 1992. On the neurobiology of vocal communication. In H. Papousek, U. Jürgens & M. Papousek (eds): *Nonverbal* *vocal communication,* pp. 31-42. Cambridge University Press, Cambridge.

Jürgens U. & Zwirner P. 1996. The role of the periaqueductal grey in limbic and neocortical vocal fold control. *Neuroreport,* 7: 2921-2923.

Juslin P. & Sloboda J. 2001. *Music and Emotion: Theory and Research*. Oxford University Press, Oxford.

Kappas A., Hess U. & Scherer K. 1991. Voice and emotion. In R. Feldman & B. Rime (eds): *Fundamentals of nonverbal behavior: studies in emotion and social interaction,* pp. 200-238. Cambridge University Press, Cambridge.

Karow C., Marquardt T. & Marshall R. 2001. Affective processing in left and right hemisphere brain-damaged subjects with and without subcortical involvement. *Aphasiology,* 15: 715-729.

Kirschner S. & Tomasello M. 2009. Joint drumming: Social context facilitates synchronization in preschool children. *J. Exp. Child Psychol.,* 102: 299–314.

Kisliuk M. 1991. *Confronting the quintessential: singing, dancing and everyday life among the Biaka Pygmies (Central African Republic)*. Ph.D. thesis, New York University.

Kitamura C., Thanavisthuth C., Burnham D. & Luksaneeyanawin S. 2002. Universality and specificity in infant-directed speech: pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behav. Dev.,* 24: 372-392.

Kochetkova V. 1978. *Paleoneurology*. Winston & Sons, Washington.

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

Koelsch S., Grossmann T., Gunter T., Hahne A. & Friederici A. 2003. Children processing music: Electric brain responses reveal musical competence and gender differences. *J. Cogn. Neurosci.,* 15: 683-693.

Kraut R. & Johnston R. 1979. Social and emotional messages of smiling: an ethological approach. *J. Pers. Soc. Psychol.,* 37: 1539-1553.

Krumhansl C. 1997 An exploratory study of musical emotion and psychophysiology. *Can. J. Exp. Psychol.,* 51: 336-352.

Krumhansl C. & Schenck D. 1997. Can dance reflect the structural and expressive qualities of music? A perceptual experiment on Balanchine's choreography of Mozart's Divertimento No. 15. *Music. Sci.,* 1: 63–85.

Kudo H. & Dunbar R. 2001. Neocortex size and social network size in primates. *An. Behav.,* 62: 711-722.

Kuhl P. 1988. Auditory perception and the evolution of speech. *Hum. Evol.,* 3: 19-43.

Laitman J. & Heimbuch R. 1982. The basicranium of Plio-Pleistocene hominids as an indicator of their upper respiratory systems. *Am. J. Phys, Anthropol.*, 59: 323-344.

Laland K. & Brown G. 2011. *Sense and Nonsense: Evolutionary Perspectives on Human Behaviour* (2nd Edition). Oxford University Press, Oxford.

Lavy M. 2001. *Emotion and the Experience of Listening to Music: A Framework for Empirical Research*. Ph.D. Thesis, University of Cambridge.

Lawson G. & d'Errico F. 2002. Microscopic, experimental and theoretical re-assessment of Upper Palaeolithic bird-bone pipes from Isturitz, France: ergonomics of design, systems of notation and the origins of musical traditions. In E. Hickman, A. Kilmer & R. Eichman (eds): *Studien zur Musikarchäologie III,* pp. 119–142. Verlag Marie Leidorf, Rahden.

Le Gonidec M.-B., Garcia L. & Caussé R. 1996. Au sujet d'un flûte Paléolithique. En souvenir du Dominique Buisson. *Antiquités Nationales,* 28: 149-152.

Levenson R., Ekman P. & Friesen V. 1990. Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology,* 27: 363-384.

Levi-Strauss C. 1970. *The Raw and the Cooked* (Translated by J. & D. Weightman). Cape, London.

Levitin D. & Bellugi U. 1997. Musical abilities in individuals with Williams Syndrome. *Paper presented at the 1997 Society for Music Perception and Cognition.* Massachusetts Institute of Technology, Cambridge MA.

Lewis J. 2009. As well as words: Congo pygmy hunting, mimicry and play. In R. Botha & C.

Knight (eds): *The Cradle of Language*, pp. 236-256. Oxford University Press, Oxford.

Lewkowicz D. J. 1998. Infants' response to the audible and visible properties of the human face: II. Discrimination of differences between singing and adult-directed speech. *Dev. Psychobiol.,* 32: 261-274.

Liégois-Chauvel C., Giraud K., Badier J.-M., Marquis P. & Chauvel P. 2003. Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of human auditory cortex. In I. Peretz & R. Zatorre (eds): *The Cognitive Neuroscience of Music*, pp. 152-167. Oxford University Press, Oxford.

Locke D. 1996. Africa: Ewe, Mande, Dagbamba, Shona and BaAka. In J. Titon (ed): *Worlds of music: an introduction to the music of the world's people* (3rd Edition), pp. 83-144. Schirmer, New York.

Locke J., Bekken K., McMinnLarson L. & Wein D. 1995. Emergent control of manual and vocal motor activity in relation to the development of speech. *Brain Lang.,* 51: 498-508.

MacLarnon A. & Hewitt G. 1999. The evolution of human speech: the role of enhanced breathing control. *Am. J. Phys. Anthropol.,* 109: 341-363.

Malloch S. & Trevarthen C. 2009a. Musicality: communicating the vitality of human life. In S. Malloch & C. Trevarthen (eds): *Communicative musicality: exploring the basis of human companionship*, pp. 1-12. Oxford University Press, Oxford.

Malloch S. & Trevarthen C. 2009b. *Communicative musicality: exploring the basis of human companionship*. Oxford University Press, Oxford.

Mang E. 2000. Intermediate vocalizations: an investigation of the boundary between speech and songs in young children's vocalisations. *Bulletin of the Council for Research in Music Education*, 147: 116-121.

Marin O. & Perry D. 1999. Neurological aspects of music perception and performance. In D. Deutsch (ed): *The Psychology of Music* (2nd ed.), pp. 653-724. Academic Press, New York.

Martínez I., Arsuaga J.-L., Quam R., Carretero J., Gracia A. & Rodriguez L. 2008. Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *J. Hum. Evol.,* 54: 118-124.

Martínez I., Rosa M., Arsuaga J.-L., Jarabo P., Quam R., Lorenzo C., Gracia A., Carratero J.-M., Bermudez de Castro J.-M. & E. Carbonell 2004. Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proc. Natl. Acad. Sci. U.S.A.,* 101: 9976-9981.

Masataka N. 2000. The role of modality and input in the earliest stages of language acquisition: studies of Japanese sign language. In C. Chamberlain, J. Morford & R. Mayberry (eds): *Language Acquisition by Eye*, pp. 3-24. Erlbaum, Mahwah.

Mayberry R. & Jaques J. 2000. Gesture production during stuttered speech: insights into the nature of gesture-speech integration. In D. McNeill (ed): *Language and Gesture*, pp. 199-214. Cambridge University Press, Cambridge.

McAllester D. 1996. North America/Native America. In J. Titon (ed): *Worlds of music: an introduction to the music of the world's people* (3rd Edition), pp. 33-82. Schirmer, New York.

McClave E. 1994. Gestural beats: the rhythm hypothesis. *J. Psycholinguistic Res.,* 23: 45-66.

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

Menon V. & Levitin D. 2005. The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage,* 26: 175-184.

Merker B. 1999. Synchronous chorusing and the origins of music. *Music. Sci., Special Issue*: 59-73.

Messinger D. & Fogel A. 1998. Give and take: the development of conventional infant gestures. *Merrill-Palmer Quarterly. Journal of Developmental Psychology,* 44: 566-590.

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

Mitchell R. & Gallaher M. 2001. Embodying music: matching music and dance in memory. *Music Percept.,* 19: 65-85.

Mithen S. 2005. *The singing Neanderthals: the origins of music, language, mind and body*. Weidenfeld & Nicholson, London.

Mohr J., Pessin M., Finkelstein S., Funkenstein H., Duncan G. & Davis K. 1978. Broca's aphasia: pathological and clinical. *Neurology,* 28: 311–324.

Molinari M., Leggio M., De Martin M., Cerasa A. & Thaut M. 2003. Neurobiology of rhythmic motor entrainment. *Ann. N. Y. Acad. Sci.,* 999: 313-321.

Morley I. 2013. *The prehistory of music: evolutionary origins and archaeology of human musicality.* Oxford University Press, Oxford.

Morley I. 2012. Hominin physiological evolution and the emergence of musical capacities. In N. Bannan (ed): *Music, Language and Human Evolution*, pp. 109-141. Oxford University Press, Oxford.

Morley I. 2003. *The evolutionary origins and archaeology of music. An investigation into the prehistory of human musical capacities and behaviours*. Ph.D. thesis, University of Cambridge. Published online at http://www.dar.cam.ac.uk/dcrr

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

Morton E. 1994. Sound symbolism and its role in non-human vertebrate communication. In L. Hinton, J. Nichols & J. Ohala (eds): *Sound Symbolism*, pp. 348-365. Cambridge University Press, Cambridge.

Morton E. 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.,* 111: 855-869.

Myers F. 1999. Pintupi-speaking Aboriginals of the Western Desert. In R. Lee & R. Daly (eds): *The Cambridge Encyclopedia of Hunters and Gatherers,* pp. 348-357. Cambridge University Press, Cambridge.

Nettiez J. 1983. Some aspects of Inuit vocal games. *Ethnomusicology,* 27: 457-475.

Nettl B. 2000. An ethnomusicologist contemplates universals in musical sounds and musical cultures. In N. Wallin B. Merker & S. Brown (eds): *The Origins of Music*, pp. 463-472. MIT Press, Cambridge MA.

Nettl B. 1992. North American Indian music. In B. Nettl, C. Capwell P., Bohlman I., Wong & T. Turino (eds): *Excursions in World Music*, pp. 260-277. Prentice Hall, Englewood Cliffs.

Neubauer S., Gunz P. & Hublin J.-J. 2009. The pattern of endocranial ontogenetic shape changes in humans. *J. Anat.,* 215: 240-255.

Nishitani N., Schürmann M., Amunts K. & Hari R. 2005. Broca's region: from action to language. *Physiology,* 20: 60–69.

Nobe S. 1996. *Cognitive rhythms, gestures, and acoustic aspects of speech*. Ph.D. thesis, University of Chicago (Cited in Mayberry & Jaques, 2000).

Nyklicek I., Thayer J. & Van Doornen L. 1997. Cardiorespiritory differentiation of musically-induced emotions. *J. Psychophysiol.,* 11: 304-321.

Ojemann G., Ojemann J., Lettich E. & Berger M. 1989. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J. Neurosurg.,* 71: 316-326.

Okanoya K. & Merker B. 2007. Neural substrates for string-context mutual segmentation: a path to human language. In C. Lyon, C. Nehaniv & A. Cangelosi (eds): *Emergence of communication and language*, pp. 421-434. Springer, London.

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

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

Panksepp J. & Trevarthen C. 2009. The neuroscience of emotion in music. In S. Malloch & C. Trevarthen (eds): *Communicative Musicality: Exploring the Basis of Human Companionship*, pp. 105-146. Oxford University Press, Oxford.

Papousek H. 1996. Musicality in infant research: biological and cultural origins of early musicality. In I. Deliege & J. Sloboda (eds): *Musical beginnings: origins and development of musical competence*, pp. 37-55. Oxford University Press, Oxford.

Patel A. 2003. Language, music, syntax and the brain. *Nat. Neurosci.,* 6: 674-681.

Patel A., Iverson J., Bregman,M. & Schultz I. 2009. Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.,* 19: 827–830.

Patel A., Peretz I., Tramo M. & Labrecque R. 1998. Processing prosodic and musical pattern: a neuropsychological investigation. *Brain Lang.,* 61: 123-144.

Peretz I. 2003. Brain specialization for music: new evidence from congenital amusia. In I. Peretz & R. Zatorre (eds): *The cognitive neuroscience of music*, pp. 192-203. Oxford University Press, Oxford.

Petrides M., Cadoret G. & Mackey S. 2005. Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature,* 435: 1235-1238.

Platel H., Price C., Wise J., Lambert J., Frackowiak R., Lechevalier B. & Eustache F. 1997. The structural components of music perception. *Brain,* 120: 229-243.

Poeppel D. & Hickok G. 2004. Towards a new functional anatomy of language. *Cognition,* 92: 1-12.

Provasi J. & Bobin-Begue A. 2003. Spontaneous motor tempo and rhythmical synchronisation in 2-1/2 and 4-year-old children. *Int. J. Behav. Dev.,* 27:220–231.

Repp B. 1991. Some Cognitive and perceptual aspects of speech and music. In J. Sundberg, L. Nord & R. Carlson (eds): *Music, Language, Speech and Brain*, pp. 257-268. MacMillan Press, Basingstoke.

Reznikoff I. 2008. Sound resonance in prehistoric times: A study of Paleolithic painted caves and rocks. *Paper presented to Acoustics Paris '08. 24th June-4th July 2008.*

Reznikoff I. & Dauvois M. 1988. La dimension sonore des grottes ornées. *Bull. Soc. Préhist. Française,* 85: 238-246.

Richman B. 2000. How music fixed "nonsense" into significant formulas: On rhythm, repetition, and meaning. In N. Wallin, B. Merker & S. Brown (eds): *The origins of music*, pp. 301-314. MIT Press, Cambrige MA.

Richter D., Waiblinger J., Rink W. & Wagner G. 2000. Thermoluminescence, electron spin resonance and C-14-dating of the Late Middle and Early Upper Palaeolithic site of Geissenklosterle Cave in southern Germany. *J. Archaeol. Sci.,* 27: 71-89.

Rodríguez L., Cabo L. & Egocheaga J. 2003. Breve nota sobre el hioides Neandertalense de Sidrón (Piloña, Asturias). In M. Aluja, A. Malgosa & R.Nogués (eds): *Antropología y Diversidad, vol. 1*, pp. 484–493. Edicions Bellaterra, Barcelona.

Sacks O. 2008. *Musicophilia: tales of music and the brain*. Vintage, London.

Schellenberg E. & Trehub S. 1996a. Natural musical intervals: evidence from infant listeners. *Psychol. Sci.,* 7: 272-277.

Schellenberg E. & Trehub S. 1996b. Children's discrimination of melodic intervals. *Dev. Psychol.,* 32: 1039-1050.

Schellenberg E. & Trehub S. 1994. Frequency ratios and the perception of tone patterns. *Psychon. Bull. Rev.,* 1: 191-201.

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

Scherer K. 1986. Vocal affect expression: a review and model for future research. *Psychol. Bull.,* 99: 143-165.

Scherer K. 1985 Vocal affect signalling: a comparative approach. *Advances in the Study of Bahaviour,* 15: 189-244.

Schön D., Magne C. & Besson M. 2004, The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology,* 41: 341-349.

Schulkin J. 2013. *Reflections on the musical mind: an evolutionary perspective*. Princeton University Press, Princeton.

Schulz G., Varga M., Jeffires K., Ludlow C. & Braun A. 2005. Functional neuroanatomy of human vocalization: An $H_2^{15}O$ PET Study. *Cereb. Cortex,* 15: 1835-1847.

Schweiger A. 1985. Harmony of the spheres and the hemispheres: the arts and hemispheric specialisation. In D. F. Benson & E. Zaidel (eds): *The dual brain,* pp. 359-373. Guildford Press, New York.

Scothern P. 1992. *The music-archaeology of the Palaeolithic within its cultural setting*. Ph.D. thesis, University of Cambridge, Cambridge.

Sergent J., Zuck E., Terriah S. & MacDonald B. 1992. Distributed neural network underlying musical sight-reading and keyboard performance. *Science,* 257: 106-109.

Shennan S. 2002. *Genes, memes and human history*. Thames & Hudson, London.

Shepard R. 1982. Geometrical approximations to the structure of musical pitch. *Psychol. Rev.,* 89: 305-333.

Sloboda J. 1998. Does music mean anything? *Music. Sci.,* 2: 21-31.

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

Sloboda J., Hermelin B. & O'Connor N. 1985. On the anatomy of the ritard: a study of timing in music. *J. Acoust. Soc. Am.,* 97: 53-67.

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

Spoor F., Wood B. & Zonneveld F. 1994. Implications of hominid labyrinthine morphology for evolution of human bipedal locomotion. *Nature,* 369: 645-648.

Spoor F. & Zonneveld F. 1998. Comparative review of the human bony labyrinth. *Yearb. Phys. Anthropol.,* 41: 211-251.

Steinbeis N. & Koelsch S. 2008. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb. Cortex,* 18: 1169-1178.

Storr A. 1992. *Music and the Mind*. Harper Collins, London.

Tartter V. 1980. Happy talk: perceptual and acoustic effects of smiling on speech. *Percept. Psychophys.,* 27: 24-27.

Tartter V. & Braun D. 1994. Hearing smiles and frowns in normal and whisper registers. *J. Acoust. Soc. Am.,* 96: 2101-2107.

Tobias P. 1987. The brain of *Homo habilis*: A new level of organization in cerebral evolution. *J. Hum. Evol.,* 16: 741-761.

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

Trainor L. & Schmidt L. 2003. Processing emotions induced by music. In I. Peretz & R. Zatorre (eds): *The Cognitive Neuroscience of Music*, pp. 310-324. Oxford University Press, Oxford.

Tramo M., Cariani P., Delgutte B. & Braida L. 2003. Neurobiology of harmony perception. In I. Peretz & R. Zatorre (eds): *The Cognitive Neuroscience of Music,* pp. 127-151. Oxford University Press, Oxford.

Trehub S. 2003. Musical predispositions in infancy: an update. In I. Peretz & R. Zatorre (eds): *The Cognitive Neuroscience of Music*, pp. 3-20. Oxford University Press, Oxford.

Trehub S., Schellenberg E. & Kamenetsky S. 1999. Infants' and adults' perception of scale structure. *J. Exp. Psychol. Hum. Percept. Perform.,* 25: 965-975.

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

Trevarthen C. 2014. Split Brains and Human Consciousness. In D. Wright (ed): *The Elsevier International Encyclopedia of the Social and Behavioral Sciences (2nd Edition)*. Elsevier (Pergamon), Amsterdam (in press).

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

Turino T. 1992. The music of Sub-Saharan Africa. In B. Nettl, C. Capwell, P. Bohlman, I. Wong & T. Turino (eds): *Excursions in World Music*, pp. 165-195. Prentice Hall, Englewood Cliffs.

Turner R. & Ioannides A. 2009. Brain, music and musicality: inferences from neuroimaging. In S. Malloch & C. Trevarthen (eds): *Communicative Musicality: Exploring the Basis of Human Companionship*, pp. 147-181. Oxford University Press, Oxford.

Vanderhorst V., Teresawa E. & Ralston H. 2001. Monosynaptic projections from the nucleus retroambiguus region to laryngeal motoneurons in the rhesus monkey. *Neuroscience,* 107: 117-125.

Wagner H. 1989. The peripheral physiological differentiation of emotions. In H. Wagner & A. Manstead (eds): *Handbook of psychophysiology: emotion and social behavior*, pp. 77-98. Wiley & Sons, Chichester.

Wallin N., Merker B., & Brown S. 2000. *The Origins of Music*. MIT Press, Cambridge MA.

Waterman C. 1991. Uneven development of African ethnomusicology. In B. Nettl & P. Bohlman (eds): *Comparative Musicology and Anthropology of Music*, pp. 169-184. University of Chicago Press, Chicago.

Watt R. & Ash R. 1998. A psychological investigation of meaning in music. *Music. Sci.,* 2: 33-54.

Werker J., Pegg J. & McLeod P. 1994. A cross-language investigation of infant preference for infant-directed communication. *Infant Behav. Dev.,* 17: 323-333.

Wild B., Erb M. & Bartels M. 2001. Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Res.,* 102: 109-124.

Wray A. 1998. Protolanguage as a holistic system for social interaction. *Lang. Commun.,* 18: 47-67.

Zatorre R. 2003. Neural specializations for tonal processing. In I. Peretz & R. Zatorre (eds): *The cognitive neuroscience of music*, pp. 231-247. Oxford University Press, Oxford.

Zatorre R., Evans A., Meyer E. & Gjedde A. 1992. Lateralization of phonetic and pitch processing in speech perception. *Science,* 256: 846-849.

Zentner M. & Eorola T. 2010. Rhythmic engagement with music in infancy. *Proc. Natl. Acad. Sci. U.S.A.,* 107: 5768–5773.

Associate Editor, Emiliano Bruner