Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Human Evolution 50 (2006) 587–593

Lead Book Review

"Music" means nothing if we don't know what it means

Steven Mithen, The Singing Neanderthals: The Origins of Music, Language, Mind and Body, Weidenfeld and Nicolson, London, 2005, ISBN 272 pp., £8.99 (paperback), £20.00 (hardcover) ISBN 0-297-64317-7.

The Singing Neanderthals is an ambitious attempt by Steven Mithen to tackle the evolution of music and language. He argues for a precursor common to both in the form of a mode of communication that is *H*olistic - lacking compositional structure and combinatoriality; *M*anipulative - concerned with affecting the behaviour of others; *M*ulti-*M*odal - involving voice and gesture; *M*usical - making use of rhythm and melody and involving synchronisation and turn-taking; and *M*imetic - imitative and involving intentional representations: in short, "Hmmmmm" (or more concisely Hm5). Mithen's approach differs from previous attempts to describe a "musilanguage" stage in human evolution (e.g., Brown, 2000) by drawing extensively on Wray's (1998) concept of a holistic proto-language. Mithen states that "the result is a complete account not only of how music and language evolved but also of how they relate to the evolution of the human mind, body and society" (p.7). Unfortunately, the book falls considerably short of this high-reaching goal. The central problem with the argument, beyond the fact that Wray's views on the evolution of language are themselves contestable, is that it lacks clear definitions of both music and language. Music is never clearly distinguished from Hm5 and is all too often located in contexts, concepts and ideas that are profoundly ethnocentric. It seems that the book would have benefited from taking a broader view of language and meaning and taking more explicitly into account models of both animal and human communication.

Mithen puts forward his views in a clearly structured book and discusses many of the topics and literature central to this area of research. In the first eight chapters he considers arguments against music being an evolutionary spandrel, adduces psychological and neuroscientific evidence for a dissociation between music and language, explores the phenomenon of talking and singing in mother-infant interactions, and reviews the use of music in psychological and physiological therapy. The central section of the book, comprising six chapters, discusses primate vocalisations and communication, communicative pressures of a savannah lifestyle, possible impacts of the move toward bipedalism, the advent of imitation, selective pressures due to sexual selection and infant altriciality, as well as sociality and group behaviour. The concluding three chapters then offer speculation on Neanderthal "musicality" and models for the transition from a musilanguage (or Hm5) stage to the use of language and music in our species.

I shall only loosely follow the structure of Mithen's argument in this review as the points I shall raise cover broad issues with relevance across several chapters. As he is primarily arguing for a greater inclusion of "music" in our evolutionary story, the focus of this discussion is on music. I start by reviewing the case, which Mithen claims to support, for music being viewed as an evolutionary adaptation. I follow this by arguing that Mithen's view of music, as no more than a relic of a previously functional communicative system that is today employed solely for hedonistic purposes, is false and stems from a narrow and ethnocentric view of music. I argue that a broad definition of music such as that presented by Ian Cross (2003a) is essential to any meaningful investigation in this area and highlights the flaws in many of Mithen's assumptions. Finally I review Mithen's proposed model for the transition from a musilanguage stage or Hm5 to modern-day language and music, highlighting the need to account for fundamental differences between models of animal (e.g., Owings and Morten, 1998) and human communication (e.g., Sperber and Wilson, 1986).

More than cheesecake?

Pinker (1997) describes music as no more than auditory cheesecake - a pseudo-communicative and non-adaptive activity that merely tickles senses and capacities that evolved as a result of other, more survival-oriented, evolutionary pressures, in particular, language. This view has been widely argued against (e.g., Cross, 1999; Dissanayake, 2000; Miller, 2000; Huron, 2001), and it is now generally accepted that "the evidence is enough to suggest that an adaptive model for music should be the default hypothesis" (R. Foley, quoted in Balter, 2004). The main arguments for this are:

- Every known human society has what trained musicologists would describe as "music" (Blacking, 1995), hence we can expect that music accompanied *Homo sapiens* out of Africa (Cross, 2003a).
- We spend a lot of time doing it [e.g., Mekranoti Indians (Werner, 1984)].
- 6-month-old infants are capable listeners (see Trehub, 2003a) and display proto-musical abilities from birth (Papousek, 1996).

doi:10.1016/j.jhevol.2005.12.003

- Peretz and colleagues argue for the existence of specialised brain structures for music. This is debatable (see below) but the existence of Amusics who retain emotive and referential linguistic abilities is suggestive (Peretz, 2003; Peretz and Coltheart, 2003).
- Music has been around for a long time. The earliest archaeological evidence for music comes from Geissenklösterle 36,000 years ago (Morley, 2003; D'Errico et al., 2003). We should expect, however, that the use of non-fossiling instruments (e.g., bamboo, wood, etc.) and the voice preceded considerably any archaeological evidence (see Morley, 2003).

The central problem with Pinker's hypothesis is that he adopts a narrow and ethnocentric view of music as little more than patterned sound experienced auditorily for hedonistic pleasure – a specialised, commodified and rather passive activity, rather than something we actively engage in that has a wide range of important and culturally constructed meanings (e.g. Cross, 2003a).

Mithen claims to be further countering Pinker's provocative hypothesis primarily on the basis that music is too different from language to be adequately explained as an evolutionary spin-off (p.26) and states that "our ultimate concern is with the evolved physical and psychological propensities that provide the capacities for language and music in our species, *Homo sapiens*" (p.11). This is certainly correct, but it is unfortunate that he makes no mention of Iain Morley's (2003) comprehensive attempt to relate the emergence of physiological and cognitive capabilities for music to the archaeological record. Mithen's argument throughout the book does, however, support Pinker's view that music is nothing more than entertainment - a relic of a previously adaptive communicative medium that has been rendered useless by the dominance and power of language. As such, Mithen's argument is open to many of the criticisms that have been directed at Pinker. Mithen favours a common pre-cursor to music and language in the form of Hm5, reasoning that with the advent of language Hm5 was no longer required but that "there remained a compulsion among modern humans to communicate with music" (e.g., in ceremony and/or as entertainment) (p.272). However, one of Mithen's fundamental assumptions seems to be that everything needed for music is present in Hm5. This conflation of identities persists throughout the book as we are asked to imagine hominids (who are, according to Mithen's argument, utilising Hm5) synchronising vocalisations in a communal *song* at FxJj50 or confirming social bonds through communal Hm5 *singing* at Atapuerca. Perhaps even more extreme is Mithen's response to Blacking's statement (1995) that the world would be a better place if we were all able to communicate unselfconsciously through music. "I agree" says Mithen: "bring back Hmmmmm!" (p.271).

Problematic definitions

Mithen follows Bruno Nettl's definition (1983) of music as "human sound communication outside the scope of language" and states that this is "perhaps as good a definition as we can get" (p.11). This is certainly not the case (see below) and, more importantly, doesn't actually get us anywhere. First of all, language remains undefined in this statement and could range from a system of abstract referentiality encoded in a grammatical system to broader perspectives including emotive vocal inflexions, expressive timings, and/or timbral cues – i.e., features that Mithen describes as being "musical". Secondly, this definition would appear to include many things we would be unlikely to consider to be music; I could, for example, easily communicate my disagreement with this definition with a sigh or a grunt that is neither linguistic nor musical. Additionally, it should be clear that linguistic interaction itself, broadly conceptualised, incorporates all features ascribed to Hm5. As such, "musical" abilities as defined above can still be explained as having evolved for language (more precisely, non-musically contextualised interaction) with music itself remaining entirely explicable as sensory cheesecake. If we move away from the specific definitions used to the ways in which Mithen describes music and language, things remain equally problematic. At various times Mithen describes music as "a universal language" (p. 14), the "language of emotion", as making "fewer cognitive demands than language" (p.23), and, in contrast to language, as being non-referential and manipulative. Although it needs to be explored more carefully in a wider context of meaning and communication (see below), the last of these is a good and important point. The others are, however, more problematic and require further discussion.

Ethnomusicologists have long argued against music being a universal language (e.g., Bohlman, 2000) in that its meaning and culturally meaningful production is dependent on cultural knowledge and immersion. In other words it is culturally constructed. Moreover, if music is functional in any way it is likely to embody meaning for those practising it. This meaning cannot exist without being socially embedded. This is not to say that we cannot enjoy listening to, or engaging with music from other cultures but we must accept that, without a cultural setting, this may be all it is - surface level enjoyment (i.e., cheesecake). The issue of meaning will be discussed in more detail below.

The populist notion of music as the language of emotion is, also, fraught with difficulties. It is undeniable that music has close connections with emotion and that engaging with music can have a powerful effect on our emotional state. In Mithen's writing though, there is again confusion about what we understand by music. He states that "emotional expression is more central to music than to language" (p.24) and uses the example of a "song-like cry" of a bereaved mother being more emotionally affecting than any words could be. However, just as with the disagreeing grunt mentioned above, this cry is not music. In Chapter Seven Mithen discusses Deryck Cooke's attempt in 1959 to identify relationships between structural elements within a musical composition and corresponding emotions. He suspects that the central idea is correct and that there "are associations universal to all humans that have yet to be fully explored" (p.90).

From the perspective of Cooke's traditional musicological analysis his suspicion is most likely misplaced or at least buried below layer upon layer of cultural construct. The principal limitation of this approach is that it ignores the fact that music's emotional affect is predominantly active, and hence functional, in social contexts. This point, also, will be developed below.

The statement that music makes fewer cognitive demands than language is, especially from an evolutionary perspective, untenable. Mithen argues that ''you can simply sit back and let the music 'wash over' your body and mind without having to concentrate or actively listen at all''. This is akin to explaining language as having evolved so that one can listen to popular radio talk shows in the bath rather than for interpersonal and interactive communication. Music must be thought of not as something we passively consume but as something that we *do*. By interacting and communicating with others in ''real time'' music requires us to understand each other as ourselves and creates a sense of joint intentionality (Cross, 2005), to provide a joint internalisable temporal framework (Bispham, 2005), to couple actions with perceptions (e.g. Ascherleben et al., 2003), and to be able to appreciate that our actions are ''about'' something and are in some form meaningful (Tolbert, 2001).

A further issue, related to the question of what music and musicality are, is that of perfect pitch (more accurately, *absolute* pitch). This issue is raised repeatedly in the book and needs clarification. Mithen puts forward the view that musical ''savants'' (p.43), infants (pp.66-67) and Neanderthals (p.234) are or were more ''musical'' than normal modern-day adults because of their maintenance of ''perfect'' pitch. This view is built upon the premise that as we develop we unlearn ''perfect'' pitch in favour of relational processing as a result of language acquisition and the consequent need to monitor direction and distances of contour changes in speech. ''Perfect'' pitch is seen as an aid to music by ''making practice more accurate and facilitating direct connections between physical movements or instrument keys and the desired pitch'' (p.78). The developmental evidence (Saffron and Griegentrog, 2001) that Mithen draws upon is contested and remains at odds with much of the literature in this area (Trehub, 2003b; Plantinga and Trainor, 2005). Even more important, though, is the fact that *having absolute pitch does not make a person absolutely musical*! Unless we assume that fine-scale instrumental tuning preceded music, we must accept that monitoring direction of pitch movement and distance are as important, if not more so, in music as in speech. When Mithen describes musical savants as having a ''seemingly complete mastery of the domain of music'' (p.44) he is most likely describing a highly systematic mind (e.g., Baron-Cohen, 1995) and a corresponding ability to create a ''correct'' sequence of pitches on a piano. This is far removed from being able to interact musically with others, establish or create meaning through music or from being musically expressive. Mithen's view on Neanderthal ''musicality'' is also called into question by these arguments and in any case cannot amount to more than speculation.

What is music?

An evolutionarily functional definition of music needs to be able to account for music as both a biological and a cultural phenomenon; to differentiate music from other communicative acts, in particular, language; and must highlight features that may have adaptively selected for music in evolution.

Cross's (2003a) definition stresses three cross-culturally key features of music: 1) *embodiment* - physiologically grounded engagement and affect; 2) *entrainment* - being and doing together in time; 3) *transposable intentionality* - meaning that is flexible and ambiguous. These are features that Mithen touches on (e.g., pp. 15, 18) but he conveys to the reader only limited understanding of the concepts and the degree to which they are important.

As far as *embodiment* is concerned, as Mithen quite rightly notes, music compels us to move (p.15). This is true to the extent that there is no dance without music (Nettl, 1983) and even to the extent that the separation of dance and music can be seen to be highly ethnocentric (Arom, 1991). In all styles of music, bodies produce music and in most (especially improvised styles), body actions are as much a determinant of the melody and rhythm produced as is any ''choice'' of pitch patterns [e.g., Afgani dutars (Baily, 1985); blues guitar (Nelson, 2002)]. Equally, choreography, whether it be directed toward sound production or not, is an essential component of performance and the creation of meaning (as in, for example, Japanese Taiko drumming). It appears that the connection between music and movement is so strong that even passive listening (cross-culturally, not the focal mode of engagement with music) activates areas of the brain associated with motor behaviour (Janata and Grafton, 2003). This connection is, possibly, due to musical (in particular rhythmic) and timed and untimed motor behaviours sharing at least a subset of internal oscillatory mechanisms[1] (Bispham, 2003). In effect, music has access to emotions at a physiologically grounded level that are either not accessible or are less accessible in other forms of communication.

Entrainment is, in short, the process that allows us to be and to act together in time. In the experience of music, internal oscillatory mechanisms attune to regularities and periodicities in external stimuli providing reliable expectations as to the timing of future events and thus maintaining a co-dependent temporal relationship (e.g. Clayton et al., 2004). Entrainment has also been argued to provide a framework for interaction and expression in mother-infant and in linguistic interactions (e.g., Webb, 1972; Auer et al., 1999; Malloch, 1999). However, degrees of regularity in music and a constant coupling of action with perception arguably endow the entraining phenomenon with a unique degree of permanence and unambiguity. While entrainment has been proposed as the root or enabling feature of human and animal temporal perception

[1] Maybe the single most persuasive piece of evidence here is that entrainment to perturbation in a pulsed stimulus occurs below threshold of awareness - feeding directly into the motor system (Stephan et al., 2002).

(Jones, 1976), sustained periodically structured synchronicity appears rare in the animal kingdom (Merker, 2000). Additionally, it seems that we may be the only species able to substantially and volitionally alter the temporal rate at which we synchronise (Repp, 2001). As with embodiment, this form of action-perception coupled entrainment can be viewed, as sharing overlapping internal oscillatory mechanisms with the motor system[2](Bispham, 2003). Through this connection, and by providing a degree of interpersonal coupling (e.g. Benzon, 2001), entrainment enables, or at least provides the foundations for a physiologically grounded group co-regulation of state. This can be envisaged by imagining collective drumming of a simple rhythm in a group, and slowly increasing or decreasing the tempo whilst further imagining the effect this has on the way you feel physically and emotionally both individually and as a group.

In respect of *transposable intentionality*, Elizabeth Tolbert (2001, p.84) explains that ''music's power is not derived solely from syntactical or semantic referents, arousals and expectancies, or from its indexical relationships to a particular cultural context, but rather through its immediacy as a performance of socio-emotional essence and embodied gesture''. Effectively, and despite lacking direct linguistic referentiality, music would seem to ''mean'' something. Music can, for example, refer to communication with the spirits/dead (Feld, 1982), signal social transformation (Blacking, 1973), describe geographic associations (Will, 2002) or cultural narratives (Slobin, 1993). The key point that Cross is making by ascribing transposable intentionality to music is that meaning in music is both culturally constructed and individually flexible (depending on, for example, experience, social status or mood). In other words, it retains a high level of ambiguity and can mean different things to different people whilst remaining collectively meaningful (Cross, 2005). Cross quotes Goehr in suggesting that ambiguity is valuable within social and political contexts, as it ''can be used to envision an alternative culture and political order while escaping the scrutiny of censor'' (Goehr, 1993). This resonates well with Turner's view (1974, 1991) that ritualistic or totemic acts occur most prevalently at moments when verbal description would highlight paradox and also explains music's cross-culturally prevalent role in ceremony and at times of political and social change. Coupled with its ability to create a co-regulation of states, music's transposable intentionality allows it to create feelings of togetherness and of shared experience in a way and in contexts that language, due largely to its referentiality, is unable to achieve.

In short, ''*music embodies, entrains and transposably intentionalises time in sound and action''* (Cross 2003a), typically ''being expressed by means of voices and instruments that articulate patterns in pitch, rhythm and timbre and involving correlated patterns of movement that may or may not be oriented towards sound production'' (taken from Cross and Morley, 2005).

A broad definition such as this, and an equally broad definition of language, are crucial to the study of music in evolution and to any attempt to untangle music and language. Mithen stresses the importance of this latter issue and explores much of the currently available evidence for a dissociation between music and language in the brain in Chapters Three to Five. It is clear, however, from the above definition of music, that some caution is required. This discussion - centred largely around musical savants and amusics - is currently limited to questions of processing abilities that may simply reflect cultural constraints on the ways in which we engage with music and language in modern-day western society and that all too often ignore the embodied nature of musical activity. The question of music and language modules in the brain is one where, as Mithen admits (p.65), much remains to be resolved. Future progress, especially from an evolutionary perspective, must be rooted in broad and functional definitions of both music and language.

Furthermore, by identifying entrainment as an essential component of music this definition implicitly negates Mithen's view that rhythm emerged simply as a consequence of changes to our motor system (Chapter, 10). The shift to bipedalism no doubt had implications for our internal timing systems as well as freeing up the arms for communicative purposes (e.g., Trevarthen, 1999/2000), but there is no reason to believe that this is in itself more ''rhythmical'' than other forms of gait or movement systems. Crucially here, we should note that rhythm implies more than simply sonic temporal patterns resulting from movement but involves, through regularities and the process of the entrainment, the creation of mutually contingent and internalisable frameworks for interaction (Bispham, 2005).

In respect of rationales for the evolution of musical behaviour, Cross's definition of music leads us quite naturally to the view that music is primarily rooted in providing a temporal framework, collective emotionality, a feeling of shared experience, and a cohesiveness to group activities, thus supporting the efficacy of ritualistic ceremonies and the creation of alternate realities and belief systems (e.g., Roederer, 1984). As such it resonates most strongly with views of music as a social bonding agent (e.g., Benzon, 2001). As discussed on pages 216-217 of Mithen's book, Freeman (2000) theorises that oxytocin release during group music-making loosens synaptic connections in which prior knowledge is held and thus ''clears the path for the acquisition of new understanding through behavioural actions that are shared with others''. This view is scientifically somewhat contentious but chimes well, in terms of the underlying idea, with the fact that music is consistently employed in achieving an altered state of consciousness (Rouget, 1985; Thayer et al., 1994; Huron, 2001), in group ceremony, and in times of social and political change (Gregory, 1997).

Other rationales for the emergence of music that have been proposed and are discussed in Mithen's book are sexual

[2] The striking evidence for this comes from studies on Parkinson's and Huntington disease patients - two disorders that heavily affect the motor system - showing that a pulsed stimulus can have dramatic and positive effect on gait parameters (Thaut et al., 1997). This is possibly due to rhythmic stimuli providing access to periodic oscillations crucial to motor control that had become unaccessible (Bispham, 2003).

selection (Chapter Twelve) and increasing altriciality in the hominid lineage (Chapter Thirteen). Mithen finds the first of these ideas ''attractive'' but admits that the case made by Miller (2000) is weak (p.180). Unfortunately, he finds no evidence to back his attraction beyond the fact that sexual selection has played a role in human evolution. I suspect that, as with Miller, the attraction comes from an ethnocentric view of music - as individual performers displaying their prowess to ''passive'' consumers - that is severely at odds with the definition of music given above. A sexual selection model for music, as it stands, has been widely and comprehensively rejected by a range of arguments including the following:

- Music rarely figures in human courtship and hardly at all outside large cultures [e.g., Aka Pygmies (Arom and Khalfa, 1998)].
- Complex and interactive vocal displays in nature occur outside the context of mating. Monogamous gibbons, engage in daily ''pant-hoot duets'' with their partner, most likely as a means towards strengthening pair bonding (Geissmann, 2000; Geissmann and Ordeldinger, 2000).
- According to Miller's model, music lacks any specificity, and any behaviour could be reduced to these terms (Brown, 2004).
- Musical activity/ability beyond western rock musicians doesn't support the sexual selection hypothesis. It may even be that more women than men are capable of sustaining ''tuneful singing'' (Trehub, 2003a).
- ''If it was sexual selection, it would be a lot more restricted. We would see it more in courtship and less in other activities. Musical ability and activity are too widespread''. (Foley, quoted in Balter, 2004).

That increasing altriciality in the hominid lineage and a corresponding need for new (or improved) forms of mother-infant communication had paramount implications in the evolution of musical and linguistic capacities is more persuasive (see Cross, 2003b). Mithen quotes Dissanayake in explaining that the induction and concordance of emotional states became increasingly important in evolution and was achieved ''by coevolution in infants and mothers of rhythmic, temporally patterned, jointly maintained communicative interactions that produced and sustained positive affect by displaying and imitating emotions of affiliation, and thereby sharing, communicating and reinforcing'' (Dissanayake, 2000). One note of caution: just as this is not language, it is also not music. Nevertheless, this is undoubtedly a crucial factor/pressure in the move from animal to human forms of communication.

From animal communication to human communication

Mithen follows Wray (1998) in describing primate communication as being holistic and manipulative (in Wray's terminology ''functional'') and argues for it being also multi-modal and ''musical'' - ''Hm4'' (Chapter Eight). He goes on to claim that anatomical changes (p.128-129) and the pressures of a savannah lifestyle (p.132), namely EQ growth (p.127) and increasing social demands (p.134-136) in, for example, *Homo habilis* and *H. rudolfensis*, led to extensions (in scope but not in design) of the hominid Hm4 repertoire. At a later stage (unspecified but possibly with *H. heidelbergensis*) mimesis is added, in Mithen's model, to the hominid communicative system, as a result of an increased reliance upon hunting and corresponding pressures towards increased communication about the natural world. Hence Hm4 became Hm5 (p.172). Finally, Mithen describes Wray's model for the evolution of language whereby holistic phrases are broken down through a process of ''segmentation'' into separate units, each of which has its own referential meaning and can then be recombined with units from other utterances - the emergence of compositionality (p.253). He goes on to explain Wray's view that segmentation arose ''from the recognition of chance associations between the phonetic segments of the holistic utterance and the objects or events to which they relate'' and that ''once recognised, these associations might then have been used in a referential fashion to create new, compositional phrases''. Mithen argues that this process occurred after 200,000 years ago and not before, due to pressures imposed by the structure of *H. sapiens* communities, where ''people began to adopt specialised economic roles and social positions'' and where ''talking with strangers'' (Wray and Grace, in press) became ''an important and pervasive aspect of social life'' (p.257-258). In this story, language at first supplemented and then came to dominate Hm5 (p.259), with music - ''a communication system specialising in the expression of emotion'' (p.267) - emerging somewhat non-functionally out of the remnants (Chapter Seventeen).

The feasibility of Wray's model is, as Mithen admits, heavily contested (e.g., Binkerton, 2003; Tallerman, 2006) but, most importantly, the scenario described above fails to address both the crucial question of how meaning is created and fundamental differences between models of animal and human communication. In the above model, the Hm4 used by primates is simply gradually expanded to incorporate more quasi-intentional ''messages'' with correlated increases in EQ, and combinatoriality is seemingly the only thing separating holistic ''messages'' from language. As such it does little to account for the crucial ability to create intentional representations (e.g., Donald, 1991, 1997), the importance of which is strangely rejected by Mithen (p.168). Furthermore the model makes little sense in light of comparisons between animal and human communication.

According to Owings and Morton's (1998) model, animal communication occurs as a result of individuals managing and assessing signals for the purpose of regulating the behaviour of others. Importantly, managers and assessors [elsewhere ''senders'' and ''receivers'' (e.g., Seyfarth and Cheney, 2003)] need not display any awareness of their own or other's intentionalities but can simply be predisposed through processes of evolution (and conditioning) to behave in certain ways. In this model signals are seen to be indicative of an animal's internal state and endow their actions with a degree of predictability through ''motivational-structural'' meaning (a correlation between the type of signal emitted and the motivational state of an individual). By stark contrast, humans communicate by

establishing joint attention, intentionality (Tomasello, 1999; Tomasello et al., in press) and representation, and humans are ultimately concerned with ''managing'' actions *and* the cognitive environment of others (Sperber and Wilson, 1986). Getting to grips with this crucial distinction and how it relates to the questions of how we communicate and create meaning is going to be an essential step in our future understanding of the evolution of language and music (Cross, 2005; Tolbert, 2005).

I do hope that *The Singing Neanderthals* will open the door to greater involvement and interest in music by archaeologists and anthropologists researching human evolution. Mithen is right to note that he as an archaeologist has access to information and evidence that John Blacking was lacking when, in 1973, he famously asked the question: How musical is man? However, his book also highlights the need for interpretations of this nature to be rooted in cross-disciplinary collaborative efforts that are securely grounded in the current musicological and ethnomusicological literatures. What is clear from reading Mithen's book and from the increasing literature related to this area of research is that music is an essential biological and cultural component of who we - *Homo sapiens* - are. No evolutionary story can be complete without it!

References

Arom, S., 1991. African Polyphony and Polyrhythm: Musical Structure and Methodology. Cambridge University Press, Cambridge.

Arom, S., Khalfa, J., 1998. Une raison en acte: pensée formelle et systématique musicale dans les sociétés de tradition orale. TOME, vol. 84/1. pp. 5–17.

Ascherleben, G., et al., 2003. Timing mechanisms in sensorimotor synchronisation. In: Prinz, W., Hommel, B., et al. (Eds.), Common Mechanisms in Perception and Action. Attention and Performance, vol. XIX. Oxford University Press, Oxford.

Auer, P., Couper-Kuhlen, E., Muller, J., 1999. Language in Time: The Rhythm and Tempo of Spoken Language. Oxford University Press, New York.

Baily, J., 1985. Music structure and human movement. In: Howell, P., Cross, I., West, R. (Eds.), Musical Structure and Cognition. Academic Press, London, pp. 237–258.

Balter, M., 2004. Seeking the key to music. Science 306, 1120–1122.

Baron-Cohen, S., 1995. Mindblindness. MIT Press, Cambridge, MA.

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

Bispham, J., 2003. An Evolutionary Perspective on the Human Skill of Interpersonal Musical Entrainment. Unpublished Dissertation submitted in recognition of MPhil in Musicology, University of Cambridge.

Bispham, J., 2005. *Homo sapiens*: The Rhythmic Species? Paper Presented at the ''Rhythm Perception and Performance Workshop''. Bilzen, Belgium.

Binkerton, D., 2003. Symbol and structure: a comprehensive framework for language evolution. In: Christiansen, M., Kirby, S. (Eds.), Language Evolution. Oxford University Press, Oxford, pp. 77–93.

Blacking, J., 1973. How Musical is Man? University of Washington Press, Seattle.

Blacking, J., 1995. Music, Culture, and Experience: Selected Papers of John Blacking. In: Reginald Byron, R. (Ed.). Chicago University Press, Chicago.

Bohlman, P., 2000. Ethnomusicology and music sociology. In: Greer, D. (Ed.), Musicology and Sister Disciplines. Oxford University Press, Oxford, pp. 288–298.

Brown, S., 2000. The ''Musilanguage'' model of music evolution. In: Wallin, N., Merker, B., Brown, S. (Eds.), The Origins of Music. MIT Press, Cambridge, Massachusetts, pp. 389–410.

Brown, S., 2004. Evolutionary models of music: from sexual selection to group selection. In: Tonneau, F., Thompson, N.S. (Eds.), Perspectives in Ethology. Behavior, Evolution and Culture, vol. 13. Plenum Publishers, New York, pp. 231–281.

Clayton, M., Sager, R., Will, U., 2004. In time with the music: the concept of entrainment and its significance for ethnomusicology. ESEM Counterpoint I.

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

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

Cross, I., 2003a. Music and biocultural evolution. In: Clayton, M., Herbert, T., Middleton, R. (Eds.), The Cultural Study of Music: A Critical Introduction. Routledge, London, pp. 19–30.

Cross, I., 2003b. Music and evolution: causes and consequences. Contemporary Music Review 22 (3), 79–89.

Cross, I., 2005. Musicality and the human capacity for culture. Paper presented at the Symposium on Evolutionary Musicology, MPI Seewiesen, Munich, Germany.

Cross, I., Morley, I., 2005. Music in evolution and evolutionary theory: the nature of the evidence. In: Malloch, S., Trevarthen, C. (Eds.), Communicative Musicality: Narratives of Expressive Gesture and Being Human. Oxford University Press, Oxford.

D'Errico, F., et al., 2003. Archaeological evidence for the emergence of language, symbolism, and music – an alternative multidisciplinary perspective. Journal of World Prehistory 17, 1–70.

Dissanayake, E., 2000. Antecedents of the temporal arts in early mother-infant interaction. In: Wallin, N., Merker, B., Brown, S. (Eds.), The Origins of Music. MIT Press, Cambridge, Massachusetts, pp. 389–410.

Donald, M., 1991. Origins of the Modern Mind. Harvard University Press, Cambridge, MA.

Donald, M., 1997. Precis of origins of the modern mind: three stages in the evolution of culture and cognition. Behavioral and Brain Sciences 16 (4), 737–791.

Feld, S., 1982. Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression. University of Pennsylvania Press, Philadelphia.

Freeman, W., 2000. A neurobiological role of music in social bonding. In: Wallin, N., Merker, B., Brown, S. (Eds.), The Origins of Music. MIT Press, Cambridge, Massachusetts, pp. 411–424.

Geissmann, T., 2000. Duet songs of the siamang, *Hylobates syndactylus*: I. Structure and organisation. Primate Report 56, 33–60.

Geissmann, T., Ordeldinger, M., 2000. The relationship between duet songs and pair bonds in siamangs *Hylobates syndactylus*. Animal Behaviour 60, 805–809.

Gregory, A., 1997. The roles of music in society: the ethnomusicological perspective. In: Hargreaves, D., North, A. (Eds.), The Social Psychology of Music. Oxford University Press, Oxford, pp. 123–140.

Goehr, L., 1993. Music has no meaning to speak of: on the politics of musical interpretation. In: Krausz, M. (Ed.), The Interpretation of Music: Philosophical Essays. Clarendon Press, Oxford, pp. 177–190.

Huron, D., 2001. Is music an evolutionary adaptation? Annals of the New York Academy of Science 930, 43–61.

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

Jones, M., 1976. Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychological Review 83, 325–335.

Malloch, S., 1999/2000. Mothers and infants and communicative musicality. Musicae Scientiae 29–58 (special issue: Rhythm, Musical Narrative, and the Origins of Human Communication).

Merker, B., 2000. Synchronous chorusing and human origins. In: Wallin, N., Merker, B., Brown, S. (Eds.), The Origins of Music. MIT Press, Cambridge, Massachusetts, pp. 315–327.

Miller, G., 2000. Evolution of human music through sexual selection. In: Wallin, N., Merker, B., Brown, S. (Eds.), The Origins of Music. MIT Press, Cambridge, Massachusetts, pp. 329–360.

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.

Nelson, S., 2002. Melodic improvisation on a twelve-bar blues model: an investigation of physical and historical aspects, and their contribution to performance. Ph.D thesis, City University London.

Nettl, B., 1983. The Study of Ethnomusicology: Twenty-Nine Issues and Concepts. University of Illinois Press, Urbana, IL.

Owings, D., Morten, E., 1998. Animal Vocal Communication: A New Approach. Cambridge University Press, Cambridge.

Papousek, H., 1996. Musicality in infancy research: biological and cultural origins of early musicality. In: Deliège, I., Sloboda, J. (Eds.), Musical Beginnings. Oxford University Press, Oxford, pp. 37–55.

Peretz, I., 2003. Brain specialization for music: new evidence from congenital amusia. In: Peretz, I., Zatorre, R. (Eds.), The Cognitive Neuroscience of Music. Oxford University Press, Oxford, pp. 192–203.

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

Pinker, S., 1997. How the Mind Works. Allen Lane, London.

Plantinga, J., Trainor, L., 2005. Memory for melody: infants use a relative pitch code. Cognition 98, 1–11.

Repp, B., 2001. Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. Journal of Experimental Psychology. Human Perception and Performance 27, 600–621.

Roederer, J., 1984. The search for a survival value of music. Music Perception 1, 350–356.

Rouget, J., 1985. Music and Trance. University of Chicago, Chicago.

Saffron, J., Griegentrog, G., 2001. Absolute pitch in infant auditory learning of tone sequences by human infants and adults. Cognition 70, 27–52.

Seyfarth, R., Cheney, D., 2003. Signalers and receivers in animal communication. Annual Review of Psychology 54, 145–173.

Slobin, M., 1993. Subcultural Sounds: Micromusics of the West. Wesleyan University Press, Hanover.

Sperber, D., Wilson, D., 1986. Relevance: Communication and Cognition. Blackwell, Oxford.

Stephan, K., et al., 2002. Conscious and subconscious sensorimotor synchronization - prefrontal cortex and the influence of awareness. Neuroimage 15, 345–352.

Tallerman, M., 2006. Did our ancestors speak a holistic proto-language? In: Carstairs-McCarthy, A. (Ed.), Special Issue of Lingua on Language Evolution. doi:10.1016/j.lingua.2005.05.004.

Thaut, et al., 1997. Rhythmic facilititation of gait training in hemiparetic stroke rehabilitation. Journal of Neurological Sciences 151, 207–212.

Thayer, R., Newman, J., McLain, T., 1994. The self-regulation of mood: strategies for changing a bad mood, raising energy, and reducing tension. Journal for Social Psychology 67, 910–925.

Tomasello, M., 1999. The human adaptation for culture. Annual Review of Anthropology 28, 509–529.

Tomasello, M., et al. Understanding and sharing intentions: the origins of cultural cognition. Behavioural and Brain Sciences, in press.

Tolbert, E., 2001. Music and meaning: an evolutionary story. Psychology of Music 29, 89–94.

Tolbert, E., 2005. Musical Meaning, Decomposability and Symbol Grounding. Paper Presented at Science and Music Group Seminar, Cambridge.

Trehub, S., 2003a. The developmental origins of musicality. Nature Neuroscience 6, 669–673.

Trehub, S., 2003b. Commentary on Saffran and Griegentrog (2001): absolute and relative pitch processing in tone learning tasks. Developmental Science 6, 46–47.

Trevarthen, C., 1999/2000. Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. Musicae Scientiae 155–215 (Special Issue).

Turner, V., 1974. Dramas, Fields and Metaphors: Symbolic Action in Human Society. Cornell University Press, Ithaca.

Turner, V., 1991. The Ritual Process. Structure and Anti-structure. Cornell University Press, Ithaca, NY.

Webb, J., 1972. Interview synchrony: an investigation of two speech rate measures in an automated standardized interview. In: Siegman, A., Pope, B. (Eds.), Studies in Dyadic Communication. Pergamon, Oxford.

Werner, D., 1984. Amazon Journey; An Anthropologist's Year among Brazil's Mekranoti Indians. Simon and Schuster, New York.

Will, U., 2002. Oral memory in Australian Aboriginal Song Performance and the Parry-Kirk debate: a cognitive ethnomusicological perspective. In: Hickmann, E., Eichmann, R. (Eds.), Proceedings of the International Study Group on Music Archaeology, 10.

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

Wray, A., Grace, G. The consequences of talking to strangers: evolutionary corollaries of socio-cultural influences on linguistic form. Lingua, in press.

John Bispham
Department of Biological Anthropology,
Centre for Music and Science,
Cambridge University, UK
E-mail address: j.bispham@human-evol.cam.ac.uk