# Music as a communicative medium

Ian Cross
Ghofur Eliot Woodruff

Centre for Music & Science
Faculty of Music
University of Cambridge
West Road
Cambridge CB3 9DP
UK

In R. Botha & C. Knight (Eds.),
*The prehistory of language* (Vol. 1, pp. 113-144).
Oxford: Oxford University Press, 2009.

**Abstract**

This chapter draws on ethnomusicological, cognitive and neuroscientific evidence in proposing that music is a communicative medium with features that are optimised for the management of situations of social uncertainty, and that music and language constitute complementary components of the human communicative toolkit. It presents a theory of meaning in music, and compares its implications with those of a recent theory of prosodic features of language.

**Key words**: music; evolution; language; ethology; prosody

## Introduction

Like language, music appears to be a universal human capacity; all cultures of which we have knowledge engage in something which, from a western perspective, seems to be music (Blacking, 1995), and all members of each culture are expected to be able to engage with music in culturally appropriate ways (Cross, 2006). Like language, music is an interactive and participatory medium (Small, 1998) that appears to constitute a communicative system (Miell et al., 2005), but one that is often understood as communicating only emotion (Juslin & Sloboda, 2001). This immediately raises a significant question; why should an apparently specialised medium for the communication of emotion have arisen in the human species? After all, language and gesture provide extremely potent media for the communication of emotion, yet we as a species have access to a third medium - music - which would appear to be quite redundant.

However, the notion that the function of music is wholly and solely to communicate emotion is called into question by much recent ethnomusicological research, which suggests that although many of the uses of music will indeed impinge on the affective states of those engaged with it, music fulfils a wide range of functions in different societies, in entertainment, ritual, healing and in the maintenance of social and natural order (see, e.g., Feld & Fox, 1994; Titon, 1996; Nettl, 2005). These considerations shift attention away from the question of *why* we have music towards the examination of *how* it is that music can fulfil such a wide range of functions, and *what* - if anything - renders it distinct from language as a communicative medium (for language also appears capable of fulfilling the functions that have been attributed here to music).

Music performs a huge array of functions across different cultures, but one very generic feature that they all appear to share is the management of social relationships, particularly in situations of social uncertainty. Such situations include, but are not limited to:

- significant life transitions for individuals, and for individuals as part of a wider community, such as adolescence to adulthood, childhood to adolescence, life to death (Blacking, 1976, on the *domba* initiation ceremonies of Venda girls at menarche; Feld, 1982, on the Kaluli *gisalo* funerary ceremony)
- circumstances where the integrity or stability of a community is perceived to be threatened or is felt to require re-affirmation (Slobin, 1993, particularly on the social dynamics of the *klezmer* revival)
- the affirmation - or even encouragement - of propitious relationships between the social order and that which sustains it, for example, the natural environment construed as an agentive and hence prospectively social force (Turnbull, 1965, on the Mbuti pygmies' use of song to mediate their relationships to their forest environment)
- the management of inter-group relationships, e.g. by constituting a mutually accessible framework for non-conflictual between-group interactions (Clendinnen, 2005, on early music and dance encounters between British sailors and Australian aboriginal inhabitants; Marrett, 2005, on the system of ceremonial reciprocity developed at Wadeye in Northern Australia), or by serving as a mechanism for consolidating within-group bonds so as to

enhance the likelihood of success in inter-group encounters (Richards, 2006, in press, on the use of music by West African child militias)

- those involved in the formation of individual, and collective, within-group identity (MacDonald et al., 2002 on music's roles in creating and projecting identities in contemporary western cultures)
- instances of personal crisis, (Sloboda et al., 2001, on the use of music for the regulation or self-regulation of affect)
- the hazards and uncertainties of the most universal of human experiences, that of caregiver-infant interaction (Trevarthen, 1999/2000)

Of course, language also plays a significant role in such situations of social uncertainty, enabling the sharing of intentions and intentionality and articulating and sustaining the contexts of social relationships. But language possesses at least one capacity that music does not share: that of expressing and communicating propositions and propositional attitudes. At the limit, language is capable of articulating - or of being interpreted as articulating -simple and complex propositions that may bear specific and unambiguous meanings. Music, whatever else its powers, is incapable of so doing; as Blacking, (1995, p237) observed, "Not only can the 'same' patterns of sound have different meanings in different societies; they can also have different meanings within the same society because of different social contexts.". In this chapter we shall be suggesting that music's inability to express unambiguous meaning underwrites its powers to manage situations of social uncertainty and exploring a framework for understanding how music appears able to sustain such polyvalent significance.

## Language and music in situations of social uncertainty

Humans are an intensely and complexly social species (Foley, 1995; Dunbar, 1996). For such a species, communication is at a premium; any individual's likelihood of survival and reproduction will depend not only on their individual capacity to deal with their physical environment in terms of threats and resources, but also on the optimisation of their capacities to engage with conspecifics (Shennan, 2002). Hence from an evolutionary perspective, at first glance language, with its capacity to denote resource location, disseminate resource availability, broadcast perception of danger, enhance transmission of behaviour, share intentions and intentionality, and sustain social relationships, appears to offer all that might be required in terms of an effective communicative medium.

But the use of language, with its potential for specificity and singularity of meaning, may also pose problems. When social situations are on the edge - encounters with strangers, changes in social affiliations, disputed courses of action - the fact that language can be interpreted as unambiguously denoting individual feelings, attitudes and intentions can tip situations over into conflict, between groups or within groups. In situations of social uncertainty language may become inefficacious or even dysfunctional However, as we have noted, it is in just these types of situation where music appears to be most likely to manifest itself. Why should music appear to be the preferred medium of communication - or at least, of interaction - in such contexts? To address this question it is first necessary to explore what might be the requirements on a communicative system optimally adapted to manage social uncertainty.

From an ethological perspective it can be argued that such a system should either be, *or be interpreted as*, capable of only honest signalling (revealing to the receiver qualities of the signaller that are relevant to the communicative situation: Szamado & Szathmáry, 2006). In other words, the process of producing signals would have to incorporate features of which the interpretative scope would necessarily be constrained by an unambiguous fit between signal characteristics (acoustical and motoric) and the costly and hard-to-fake motivational-intentional and bodily states inferable as underlying signal production. The sense of an honest signal would be further enhanced were it to be rooted in mechanisms that would increase the likelihood that participants interacting in this communicative medium would experience each others' states and intentions as mutually manifest. At the same time, it would be necessary that the interpretative scope of the communicative system should be broad enough so as not to require precise alignment between the motivations and intentional states of participants. In other words, the signals produced should be polysemic or ambiguous, capable of multiple interpretations. While some of these requirements appear to be contradictory (how can a signal be honest yet polysemic?), it will be argued here that music provides an example of a communicative medium that conforms to all these requirements.

## Music as 'honest signal'

The idea that music constitutes an 'honest signal' - or at least, a 'natural sign' is embedded in much thinking about music. The widespread and longstanding view (traceable back to classical Greek thought) that music is primarily a communicator of emotion has already been alluded to, and is not unique to western conceptions of music (see, e.g., Basso, 1985). Thinking about music from an evolutionary perspective, Darwin (1872/1998) was happy to adopt the 'remarks' of 'Mr Litchfield' (one of his correspondents, who had 'long attended to the subject of music') in suggesting that music mirrors or captures the relationships between affective state and sound that are found across a wide range of species, embodying in the musical signal clues as to the emotional state of its producer. As he puts it (ibid., p94), "A great part of the emotional effect of a song depends on the character of the action by which the sounds are produced", that 'character' being contingent on the emotional state of the music's producer (though Darwin is careful to note that "this leaves unexplained the more subtle and more specific effect that we might call the *musical* expression of the song…" [italics in original]). Hence music, as an expression of emotion, constitutes an 'honest signal' in revealing to a listener qualities of the music's producer that are necessarily concomitant on the nature of the signal.

Music certainly appears to be experienced as having consistent, though very broad, emotional significance for listeners, a fact that has been exploited in numerous experiments that have explored the behavioural, cognitive, and neural underpinnings of affective states and have used music to elicit generically different affective states as a highly effective Mood Induction Procedure (e.g., Albersnagel, 1988). Music is also a ubiquitous facet of the experience of film and other media such as computer games, being employed to manipulate the emotional responses of audiences to the unfolding cinematic narrative or game situation (Gorbman, 1987; Kassabian, 2001). It is notable that music tends to be used in Mood Induction Procedures (MIPs) simply on the grounds that it works, and works more effectively than other methods (such as the linguistic and biographically-based Velten MIP); rarely is any compelling rationale advanced in the Mood Induction literature for why music might have such powers. Equally, the affective powers of music in cinema and multimedia have rarely been subjected to exploration and explanation (amongst the few exceptions being the work of Boltz, 2001; 2004, and Cohen, 2001). However, recent theories of animal vocal communication might afford some clues as to music's efficacy in influencing the affective states of listeners.

In a recent review of animal communication systems, Seyfarth and Cheney (2003) conclude that most, if not all, non-human animal communication is achieved almost contingently, as signallers are unaware of the means whereby their vocalisations effect their consequences. Yet a substantial literature (see, e.g., Owren & Rendall, 2001) suggests that there are at least some consistent correspondences between signal structure and signal outcome in animal communication. These consistencies form the focus of Owings and Morton's (1998) approach, which aims to remove some of the contingency from animal acoustic interaction by dispensing with the notion of information transfer and adopting a pragmatic approach, suggesting that animals, in producing vocal and other sounds, are generally seeking to manage their physical and social environments rather than to transmit information. At the same time, other animals – conspecifics and members of other species – are seeking to assess the import of the signals emitted by others. This 'management-assessment' framework is largely governed by relationships between the acoustical characteristics of the signals produced and the biological consequences that can be inferred from the acoustical features of the signal.

Processes operating over evolutionary timescales act to integrate sensitivities to those relationships into the sensory and behavioural repertoires of animals. These sensitivities can then be brought to bear in the inferences made in assessment of prospective biological consequences associated with the signal characteristics as well as in the use of signals with particular features to manage the social or physical environment. A simple example would be the case of a small animal seeking to defend a territory in an environment where little visual information is available, who might produce signals at as low a frequency as can be achieved to broadcast an impression of large size. Were this to be achievable at a lower cost to the sound producer than the cost likely to be incurred by engaging in physical combat, then the use of such signals - and the capacity to produce ever lower-frequency signals - would afford a strong evolutionary payoff (in terms of survival and reproduction) for those members of the species more capable of producing low-frequency signals consistently at a low cost to themselves, who would then be likely to come to dominate the species' genotype. At the same time there would also be selection for those members of the species most capable of correctly inferring likely size (and hence, likelihood that the sound producer constitutes a real and prospective threat, capable to inflicting physical damage) from the sounds produced by conspecifics. Hence processes of both management and assessment would work together to generate an evolutionarily stable strategy, referred to by Owings and Morton as Expressive Size Symbolism.

These correspondences of sound structure and biological significance are not limited to the frequency domain: Owings and Morton note that, as well as frequency, other acoustical parameters of the signal, including bandwidth and intensity, tend to covary with the prospective significance of biological situations. Evolutionary processes act to inscribe in the genomes of many species predispositions towards particular motivational states according to different features of the acoustical signals they encounter and produce in the form of what Owings and Morton term a "motivational-structural" code. They are thus postulating a close relationship between the motivational states of organisms (governed by affective systems) and the global structural characteristics of acoustic signals.

It would be surprising if such motivational-structural principles did not account for significant aspects of the human response to sound, given the extent to which we share appetitive, reproductive, sensory and limbic systems with many other species; the dynamics of at least the latter two are likely to have been profoundly shaped by the history of our predecessor species' interaction with the regularities of the environments they encountered and successfully navigated. Hence a close relationship can be postulated between the motivational states of listeners and the global structural characteristics of musical sound. Indeed, while the now substantial experimental literature on affective responses to music (for a review, see Juslin & Sloboda, 2001) finds very few consistencies in listeners' affective responses to music, those few that can be found seem to relate *either* to whether or not the music employed is personally selected by, and hence is meaningful to, the participants, *or* to global structural characteristics of the music (Evers & Suhr, 2000; Husain et al., 2002). Moreover, the idea that music may incorporate features that can be interpreted in motivational-structural terms is supported by findings that music has consistent effects on animal behaviours that are interpretable as a consequence of the capacity of its global structural characteristics to 'modify physiological arousal levels' (Rickard et al., 2005, p252).

One can think of motivational-structural principles as constituting a locus for natural meaning in music, endowing music with the potential to be employed and interpreted as an honest signal. Nevertheless, experimental evidence (e.g., Blood & Zatorre, 2001) suggests that listeners' affective responses to music are also mediated by other factors, notably prior personal engagement with a particular piece or genre. This account of the operation of motivational-structural processes has parallels in the work of Pavlov, who noted that even when a stimulus-response relationship could be classified as *unconditioned* or *innate* it could be over-ridden by conditioning. Hence the actuation of motivational-structural principles in listeners' responses to music does not mean that the significance for listeners of any music that activates these principles is fixed. While roughly consistent motivational states may be elicited by a piece of music in listeners, this condition is not sufficient unambiguously to endow the piece with a specific meaning. However, the operation of motivational-structural processes is likely to set limits on the range of possible significances that may be abstracted by listeners from a given musical stimulus; motivational-structural principles hold the experience of meaning in music on a leash.

## The human dimensions of musicality

Hence motivational-structural principles are best thought of as contributing only one aspect or dimension of music's capacity to signify; other factors, such as prior personal associations, must be adduced that endow it with a sense of communicable meaning. Responses to music are evidently motivated by a history of personal engagement (Davies, 1978), yet this history is not wholly individual and personal; it is mediated by, and rooted in, culture. Responses to, and indeed, capacities for, music are the result of active participation in, and engagement with, the dynamics and specificities of particular cultural contexts and processes, as well as of individual life histories. They are shaped by the conceptions and uses of music that exist within a specific cultural framework (Nettl, 2005), by the contingencies of cultural formation and change (Feld, 1996), by enculturative, formal and personal learning processes (Deliège & Sloboda, 1996), and by associations of music with episodes in and aspects of an individual's life history (MacDonald et al, 2002).

The consequences of these participatory and culturally-specific factors and processes can be thought of as providing the substrate for a culturally-enactive dimension to musical meaning, co-existing with the motivational-structural dimension. In contrast to the principled nature of the latter dimension, the operations of the culturally-enactive dimension in attributing meaning to music may appear intensely arbitrary. They are evidenced, for example, in the use of music as cultural emblem (as in the use of different pieces of music by the followers of different football clubs); or in the categorical distinctions that might be drawn within a culture between music and other phenomena - often religious or liminal - that, from the perspective of another culture, might appear to be musical; or in the existence of constraints on the types of musical behaviours held to be appropriate for different age-groups in certain societies.

We can now postulate two dimensions in the experience of meaning in music, one which relates to aspects of our experience of the world that are conditioned by our biological heritage and that may have some cross-species generality - the **motivational-structural** dimension - and one which derives from the particularities of the cultural contexts in which we develop and come to play a part - the **culturally-enactive** dimension. Both dimensions would be simultaneously operational in the experience of music, yielding a *sense* of an honest signal but allowing meanings in music to appear (within personally and culturally defined limits) fluid and contestable. This distinction between the biological and the cultural in articulating the dimensions of musical experience appears neatly to conform to the venerable nature-nurture dichotomy. However, there are aspects of the experience of musical meaning that call such a simple dichotomous account into question.

Even in the absence of cultural knowledge we can experience the sound of the music of another culture **as** music; when we encounter the sounds of Japanese Noh theatre, or of Banda initiation music from Central Africa performed in hocket (each member of the group playing alternate single notes) on *ongo* (bark trumpets), we are likely to try to bring to bear on that experience the cognitive resources appropriate to the experience of music within our own culture, with greater or lesser success. We may profoundly misinterpret the other culture's music, but we are likely to experience it as music. While motivational-structural principles may play a role in shaping such experiences, other factors would seem to be implicated. We seem to experience the music of another culture as music in ways that are different from those in terms of which we might experience other sets of sounds (such as the contingent sounds of the natural world) as music. In considering how we make sense of 'other' musics, John Blacking suggested that (1995, p238), "there must be supra-cultural cognitive resonance, and that there must be levels at which different composers, listeners and musical systems use the ‘same’ musical modes of thought.". It can be suggested that the locus of such 'supra-cultural cognitive resonances' is to be found in the similarities in the ways in which music embodies, and is the result of, particularly *human* modes of interaction across cultures, deriving from a generic human 'capacity for culture'.

Recent explorations of the human capacity for culture have focused on the nature of 'theory of mind' (see Gopnik, 1999): the ways in which animals may attribute mental states to others. It appears that only humans can be conceived of as possessing a full-blown 'theory of mind': as Call and Tomasello (2005, p261) note, even our nearest primate relatives, chimpanzees, "have the cognitive skills to recall, represent, categorize and reason about the behaviour and perception of others, but not about their intentional or mental states". In a recent paper Tomasello et al. (2005, p680) have proposed that the human capacity for culture is rooted in a capacity for, and motivation towards, 'shared intentionality' which "refers to collaborative interactions in which participants have a shared goal (shared commitment) and co-ordinated [and mutually understood] action roles for pursuing that shared goal.".

In the context of such a capacity, human communication systems would have to be understood as motivated (at least in part) by a need to make inferences about others' states of mind and intentions as well as about the physical and biological contexts of any communicative act. Thus human communicative signals must be understood as conditioned as much by the need to establish a common cognitive context for the act of communication as they are by the need to communicate information (see, e.g., Sperber & Wilson, 1995). This particular aspects of human communicative systems can be understood as underlying a dimension of musical meaning that enables Blacking's 'supra-cultural cognitive resonances' in the experience of music, and can be referred to as music's socio-intentional dimension.

This dimension would be oriented towards attributions and interpretations of human agency and intentionality in engagement with music (see also Watt & Ash, 1998). It would be rooted in performative actions and sound structures that could be interpreted as affording cues about shared intentionality that direct attention in interaction. These actions and sound structures could be interpreted as declarative or imperative, concerned with the direction of another's attention to an object or event distinct from the individuals involved in the interaction. They may construed as disclosural or dissimulative, evoking a sense that they denote distinct and different communicative intentions. These interpretations are likely to arise because music typically exhibits structural characteristics that are directly analogous to features manifested in speech and that are of significance in establishing the *pragmatic* contexts of utterances (Wilson and Wharton, 2006). Prosodic characteristics of speech such as intonation, rhythm and stress help partition the discourse into meaningful articulatory units. These units serve not only to reinforce syntactic and semantic structures but also to signal the communicative intentions of the speaker. These intentions may be manifest even in situations where propositional content is absent, such as when listening to a foreign speaker convey a message without understanding precisely what is being said. Moreover, such features are realised not only in the acoustic domain but also through physical gestures (Kendon, 2004; Krahmer & Swerts, 2007).

The socio-intentional dimension of music can be thought of as fundamentally pragmatic in relying on gestural and acoustical cues to impart a sense of communicative intent. It is likely to relate not so much to *what* unfolds musically as *how* the music unfolds, being experiential correlates of the music's temporal structure and bound up with processes of expectation and anticipation (see, e.g., Huron, 2006) or inhering in features of the musical surface such as melodic contour. The operation of this dimension of musical meaning is unproblematically evident in contexts in which music involves interactive participation. Here, music can be thought of as exhibiting features similar to those of linguistic dialogue, being exemplified in specific contoural and accentual structures (Palmer & Hutchins, 2006), call and response patterns or antecedent-consequent phrase structures. The socio-intentional domain would also be present even in contexts that involve apparently passive listening to recorded music. In such contexts it would be experienced in terms of traces of human behaviour, embodying cues as to human action or intention (or to the body-imagistic schemas that may underlie human action and interaction) in ways that have the semantic open-ness to afford the experience of joint action, joint attention and joint intention.

The three dimensions of musical meaning postulated above - the motivational-structural, culturally-enactive, and socio-intentional - are all likely to be co-present in any experience of music or engagement in musical behaviour. Music is a medium that rests on semantic indeterminacy, which has elsewhere been referred to as 'floating intentionality' (Cross, 2003). Engagement with music thus affords access to multiple and simultaneously available layers of meaning, allowing participants in a musical behaviour to interpret the significance of the music individually and independently while collectively affording to participants a sense that the music embodies an honest signal. Hence music can be viewed as embodying the characteristics of the medium optimally adapted for the management of social uncertainty outlined earlier, though one feature attributed to that medium, that of being rooted in mechanisms that would increase the likelihood that participants would experience each others' states and intentions as mutually manifest, remains to be properly addressed for the case of music.

Across, cultures, music is typically experienced as structured regularly in time, even when the events of the musical surface do not exhibit overt temporal regularities. Engagement with music typically involves the entrainment of action and attention to a commonly inferred, more-or-less regular and periodic, pulse (Drake & Bertrand, 2001). Participants in a musical activity will regulate the temporal alignment of their musical behaviours by engaging in continual processes of mutual adjustment of the timing of actions and sounds; even those who appear engaged in 'passive listening' to music will be modulating their attention according to the ways that the flow of the music affords scope for a regular pulse to be abstracted Jones & Boltz, 1989). The ability of interacting humans to produce and to entrain to a more-or-less regular auditory stimulus through processes of mutual adjustment of period and phase (Repp, 2006) seems to underpin not only musical behaviours but a wide range of human communicative acts; recent work has shown that temporal entrainment is implicated in a variety of communicative situations and is manifested in a variety of forms, from eye movements (Richardson et al., 2006) to postural changes (Shockley et al., 2003). However, only in music will entrainment be consistently and continuously oriented around a regular pulse (Bispham, 2006).

In the context of collective musical behaviours, processes of entrainment are likely to endow the communal activity with a powerful sense of joint and co-ordinated action, allowing the emergence of a sense that aims are shared and enhancing the likelihood that participants will experience each others' states and intentions as mutually manifest. Hence entrainment processes in music provide a potent means of promoting a sense of joint affiliation that helps maintain the collective integrity of a musical act even though music's floating intentionality affords each participant the possibility of interpreting its significance quite differently. Music's semantic indeterminacy (rooted in the simultaneous availability of at least three dimensions of musical meaning), together with its affiliative powers (rooted in processes of entrainment and in its exploitation of motivational-structural principles) render it effective as a communicative medium that is optimised for the management of situations of social uncertainty.

## **Language and music as communicative systems; similarities and differences**

This framework for understanding music suggests that it can be clearly differentiated from language as a mode of communicative behaviour. In contrast to the fluidity and instability of meaning that are imputed to music in this chapter, language has tended to be identified as a communicative medium that bears explicit meaning, being capable of expressing semantically decomposable propositions (see, e.g., Hockett, 1960; Montague, 1974). Yet language can also possess many of the capacities of music; language can, self-evidently, be employed in the management of social relations. And language is often quite as ambiguous as music is here claimed to be, from the vaporous mendacity of the corporate mission statement to the significatively-freighted elusiveness of poetry (where it also exhibits many of the temporal regularities that underlie entrainment processes in music). We argue that language and music should properly be distinguished as tending towards opposite poles on a continuum of capacity for *specificity* of meaning (Cross, 2005), and that there are common frameworks applicable to the understanding of at least some meaning processes in both language and music.

We see those frameworks as most pertinent to understanding the parallels between pragmatic aspects of language in discourse and the socio-intentional dimension of the experience of music. The idea that musical meaning has a socio-intentional dimension is predicated in part on the notion that performers utilise melodic and other musical structures as a means of making manifest communicative intentions or attitudinal stances, and that listeners interpret these intentions and stances. These structures can be thought of as having a discourse function similar to prosodic structures in speech: they provide contextual and narrative cues within an interactive framework. This similarity can be seen to arise from a common origin in the human voice. The mechanisms underlying vocal production lend themselves to certain prosodic forms with communicative affordances that are equally suited to both music and language. The proposition that music and language are related in this way is supported by the work of Ohala and Gussenhoven on the discourse functions of prosodic structures in language. They outline three biological speech codes - of frequency, effort and production - to explain the widespread appearance of similar prosodic structures across cultures (Gussenhoven, 2002, Gussenhoven, 2005; Ohala, 1984; Ohala, 1994).

The frequency code is so named owing to the recruitment of the frequency dimension in the communication of basic social and power relations. The code was adapted by Ohala from Morton's original formulation of motivational structures (Morton 1977), and posits that similar principles underlying animal communication are evident within the prosody of human speech. Operating on the principle that low sounds convey large size, a speaker might utilise the lower range of their voice in conversation to convey those attitudes that have a natural correlation with large size, such as confidence, dominance, or aggression. Conversely the higher register may be used to express submission, subordination and absence of threat. Gender identity and gender roles may also be expressed through the dimension of frequency, insofar as gender role is relevant to a given community.[1] Finally, the higher vocal register is sometimes used to convey uncertainty, which is said to underscore the questioning effect of rising terminals often present in interrogatives (Bolinger, 1978). Hence, in addition to our species-specific capacity for language and propositional meaning, we communicate certain attitudinal and affective states through tone of voice.

The effort code correlates the energy expended in the production of speech with a number of states naturally associated with speaker effort (Gussenhoven, 2005). Wider prosodic contours require more effort to produce and can be interpreted as conveying enthusiasm, obligingness, authority, and insistence. Conversely, a lack of commitment, enthusiasm, or interest may be signalled by narrow contours associated with less effort. The prosodic function of *focus*, in which a word is highlighted by way of raised pitch (House, 2006; Sperber & Wilson, 1995) is also attributed to the perception of speaker effort through this code. A third code, the production code, derives from the acoustic relationship between frequency and subglottal breath pressure (Gussenhoven, 2005: see Titze, 1994, for a detailed discussion of the physiological mechanics involved). Subglottal pressure is highest at the beginning of an utterance and drops gradually as breath is exhaled, resulting in a progressive lowering of pitch known as declination (Ladd, 1984; Wennerstrom, 2001). Hence a natural, biological process links the beginnings of utterances with higher pitch and the endings with lower pitch. The production code serves to partition speech into meaningful articulatory units by exploiting this natural association: high phrase beginnings signify new topics and low phrase endings signify closure. Conversely, low phrase beginnings and high phrase endings signify continuation.

In spite of their roots in Morton's original formulation of motivational-structural principles, Ohala's and Gussenhoven's codes can be conceived as operating, both in language and in music, not in motivational-structural but in socio-intentional terms. From this perspective, musical contours are no longer construed as autonomous auditory events whose sole purpose is to elicit an aesthetic end-state. By varying the width of intervals in a melodic contour, performers can exploit the effects of the effort code to convey affective states and attitudinal stances. Wider intervals might be used to make a performer's enthusiasm or assertiveness manifest, whilst narrower

[11] The effects of the frequency code are relative to an individual's voice range: it is not suggested that all higher pitched voices sound submissive and that all lower voices are aggressive.

intervals might convey more subdued emotions or even a lack of interest. Declarative and dissimulative intentions might equally be revealed through wide and narrow intervals respectively. Peaked contours might serve to highlight ostensively certain features of a musical utterance, a function analogous to that of focus in speech prosody. In terms of the production code, ascending and descending contours can be used in performance to partition the musical discourse by providing articulatory cues at phrase boundaries. A descending terminal contour may signify closure in music by invoking the association between declination and phrase endings. Conversely, an ascending terminal contour might signal continuation. While the effects of the production code in music are still being explored empirically, there is telling anecdotal evidence that performers exploit this code. For instance, the pervasiveness of the arch contour in Western folk song (Huron, 1996), could be accounted for in terms of the efficacy of the descending terminal to indicate closure. Evidence from neuroscience suggests that the perception of phrase boundaries in music elicits similar neural responses to those of speech (Knösche et al., 2005).

The role of the frequency code in music perception has received some empirical attention (Huron et al., 2006) although much remains to be explored. That the frequency code has been implicated in the human communicative system would suggest that a sensitivity to socially relevant vocal meaning is part of our evolutionary heritage; sound at a basic social level does not constitute a *tabula rasa* on which meaning is freely inscribed as it has been endowed with social significance through evolutionary processes. However, it is not suggested that music’s interpretation is bound to these codes. The cross-cultural variability of musical forms and meanings (referred to above as rooted in music's culturally enactive dimension) is in accord with Gussenhoven's (2005) claim that culturally specific encoding can and does override the communicative functions implicit in his proposed linguistic codes.

Music and speech share other features that govern their production and perceptual form. They are bound by the temporal constraints of working memory on the integration of serial event structure in perception and production. These limitations are overcome, in part, by employing hierarchical structures to coordinate the production and perception of simultaneous organisational levels. Hierarchical structures optimise transmission and perception by enabling several levels to be accessed simultaneously, such structures being evident in the prosodic domain of speech in the grouping of hierarchically related levels of foot, word and intonational phrase (Yallop and Clark, 1990). Music has been interpreted as exhibiting similar types of hierarchical groupings; Lerdahl & Jackendoff (198, 314-332) note that parallel types of formalisms appear applicable to the prosodic domain of language and to the time-span domain of tonal music. That language and music overcome the same temporal-acoustic limitations using hierarchical modes of transmission suggests that common communicative operations are in evidence.

Metric structures also provide a temporal framework that facilitates social interaction through entrainment. The metric hierarchy of music takes the form of the interaction of isochronous beats at different temporal levels; in language, the timing of syllables or the alternation of stressed and unstressed syllables (Auer et al., 1999). Although language is less intuitively associated with such metric processes than is music, this distinction is one of degree rather than kind. That isochronous beat structures in music provide a predictive framework with which listeners align their attention is well established (Jones and Boltz, 1989). With respect to entrainment, the metric structures of music must be sufficiently quantised to facilitate the predictive mechanisms necessary for performers to engage physically with music - to play along with each other in real time. Although laboratory experiments have established a similar claim for language - that rhythmic structures serve to align speakers' and listeners' attention (Dilley & McAuley, 2007, in revision; Pitt & Samuel, 1990) - it remains unclear whether everyday speech exhibits isochrony to the degree required for physical entrainment. Speakers of English exhibit periods of isochrony, both within an individual's utterance and sustained from speaker to speaker (Couper-Kuhlen, 1993). However, such isochrony is rarely sustained or quantised to the same degree as music. Whilst studies have suggested that listeners and speakers entrain gross and fine motor movements to speech (Condon, 1986; Wilson & Wilson, 2005), this appears intermittent and is far removed from the sustained participatory engagement between individuals afforded by music. This highlights an important difference between the social function of language and music: music, like language, uses entrainment to coordinate interaction, but in music this serves a primary function and hence its cues for interaction are more strongly evident.

## Music and language in evolution

Language and music appear to share more significant features as communicative systems than differentiate them. We are arguing that what appears to distinguish them most clearly is their efficacy in different social contexts (see also Cross, 2006); both language and music are instrumental in achieving social goals, but language organises social action while music organises sociality. Language has the capacity to express unambiguous, semantically decomposable, propositions; music lacks this capacity, displaying a consistent ambiguity. Nevertheless, both language and music exemplify *symbolic* behaviours, most evident in music in the culturally-enactive dimension. Both music and language rely on codes that relate the structures of the sounds that they employ to physical - biological - causes, but music embodies features that are shared with other systems of animal communication in the form of motivational-structural principles, whereas language's prosodic codes are best conceived of as bound by processes specific to human communication and located in the socio-intentional domain. Both music and language exhibit periodicity, but language's periodicities tend towards the intermittent and, in interaction, afford reciprocity of engagement, whereas music's periodicities enable collective synchronous engagement.

Language seems closest to music when it has a phatic function (as in the exchange of conventional phrases of greeting used as a preliminary to establish social relationships). Music, in our view, exemplifies and emphasises the phatic dimension of social interaction, constituting a foundational medium for 'phatic communion' (Malinowski, 1974). This could be taken to suggest that music preceded language in the evolution of human communicative systems (as Mithen's (2005) reliance on Wray's (1998) theory of holistic protolanguage would imply). Moreover, music, in the present view, incorporates features of non-human animal communication systems in the form of the motivational-structural dimension, suggesting a degree of evolutionary continuity that does not appear so evident for language.

However, the fact that music exploits the motivational-structural dimension does not necessarily mean that 'music' should be conceived of as more ancient than language. Were music's efficacy to be located solely in the motivational-structural dimension, there would indeed be every reason to conceive of it as evolutionarily prior to language. However, music, as discussed here, relies not only on the motivational-structural dimension but also on the socio-intentional and culturally-enactive dimensions - in addition to the capacity for entrainment - for its efficacy. Moreover, it has been postulated that aspects of language rely, just as do aspects of music, on the motivational-structural code (as in Ohala's original proposal for the frequency code), although, as noted above, it is more appropriate to conceive of the operation of the frequency code in speech as situated within the socio-intentional domain. Hence the fact that there are apparent continuities between aspects of musical experience in the contemporary world and aspects of animal communicative systems does not necessarily mean that 'music' preceded language in evolutionary terms. We would argue that music and language are complementary aspects of the modern human communicative toolkit, each functioning to achieve ends in respect of which the other may be less efficacious; they are best thought of as having co-evolved, most probably appearing as discretely identifiable suites of behaviour with modern *Homo sapiens* though likely to have emerged from precursor communicative systems that embodied features of both.

Unambiguous evidence for music appears early in the modern human archaeological record in Europe in the form of the Geissenklosterle pipe, dating from about 38 kya (Conard & Bolus, 2003) and the remarkable series of pipes from Isturitz, extending from the Aurignacian to the Magdalenian (D'Errico et al, 2003). All these musical artefacts are extremely sophisticated, exhibiting many features of historic wind instruments; they are highly unlikely to be the earliest manifestations of human musicality, and it is to be expected that earlier periods may yet yield archaeological evidence for musical behaviours. Moreover the prevalence of music in native American and Australian societies in forms that are not directly relatable to recent historic Eurasian or African musics is a potent indicator that modern humans brought musicality with them out of Africa. This is not to suggest that musicality emerged, fully-fledged, with modern humans. While there is no archaeological evidence that Neanderthals possessed a faculty for musicality (the much-touted Divje Babe 'flute' has been securely shown to be the result of carnivore activity, see, e.g., D'Errico & Villa, 1997), there is equally no evidence that they did not. However, the paucity of evidence for Neanderthal symbolic behaviours could be taken to suggest that music may have played a less significant - or less effective - role for Neanderthals than it does for modern humans.

Music has a potent proximate function in the management of social relationships in situations of social uncertainty. This suggests that the possession of a capacity for musicality can be interpreted as having played a significant role in the evolutionary processes that resulted in the emergence of modern humans, in facilitating our extraordinary social and cognitive flexibility, and as continuing to play that role in consolidating and sustaining those flexibilities (Cross, 2005). However, to propose that music is likely to have had adaptive value in human evolution leaves unaddressed the question of how and why a capacity for something like musicality arose in the first place. It can be suggested that a faculty for music emerges as an exaptive consequence of progressive altriciality in the later hominin lineage as a means of co-opting and regulating the exploratory value of childhood modes of thought and behaviour into the adult repertoire (Cross, 2003, 2005).

## Conclusions, and known unknowns

The view of music presented here is, as far as we can judge, consonant with the available evidence from cognitive, neuroscientific, cross-specific and ethnographic sources. Nevertheless, many aspects remain either unresolved or require further exploration. The instantiation of motivational-structural principles in our responses to music is hinted at by the results of studies of musical affect (Husain et al., 2002; Huron et al., 2006), but requires to be evaluated more broadly, and, indeed, cross-culturally. The same issue arises in respect of the socio-intentional domain; while empirical research suggests that forms of human interaction are encoded in musical structure (e.g., Watt & Ash, 1998) and many theories of music's meanings implicitly or explicitly make such a claim (e.g., Cook, 1998; Cox, 2001), more substantive experimental research is required. Indeed, such research will need to take account of Gussenhoven's formulations and should help elucidate relationships between aspects of prosodic structure in language and their analogues in music.

A further area that urgently requires more research is that of the potential species-specificity of the human entrainment capacity. At present, there is no good evidence that non-human mammalian species can either exhibit spontaneous capacities to entrain that are comparable to human capacities (which involve intentional alignment of periodic behaviours by means of both period and phase correction) or are motivated so to do (Bispham, 2006; though see Fitch, 2006a). Until reliable data are available, the notion that entrainment is a significant and species-specific - but species-general - feature of human communicative interaction must remain hypothetical, though highly probable. Moreover, the nature of interactive temporal coupling in language remains sparsely explored; further exploration is required to shed light on the communicative affordances offered by entrainment in linguistic interaction.

It has been proposed that the faculty of language is uniquely differentiated from other human communicative modes by processes of recursion (Hauser et al., 2002), though this proposal has been challenged by researchers who claim that recursion is evident in a range of human behaviours and cognitive processes (e.g., Arbib, 2005). While many theories of music rest on the premise that it embodies recursion - or, at least, hierarchicality (most comprehensively, Lerdahl & Jackendoff, 1983) - there is a paucity of empirical evidence for this contention. Moreover, theories of music that have addressed the issue of recursion have generally done so in respect of a limited range of possible musics, generally confining themselves to western tonal music of the common-practice period from ca 1600 to 1900 (though see Hughes, 1991). Cross-cultural studies are required in order to fill this lacuna (Stevens, 2004).

While archaeology has yielded rich data concerning the early appearance of musicality in the modern human record, much remains unexplored. In particular, no early artefacts that can be unambiguously identified as musical have been found outside Eurasia (Morley, 2003), though significant Neolithic finds have been made in China (Zhang et al, 1999). It is possible that no such artefacts exist to be discovered outside Eurasia, but this seems highly unlikely given the current universality of musical behaviours across cultures; it might, however, be the case that such artefacts indeed existed but have not survived in the archaeological record. It seems more likely that such artefacts indeed exist but remain either to be discovered or to be identified; the increase in interest in, and increase in sophistication concerning, possible archaeological traces of early musicality (see, e.g., D'Errico et al, 2003) make it increasingly likely that much more will be learned about the emergence of musicality in the future.

Finally, there exists a significant empirical obstacle to our understanding of the relationships between language and music. While a vast body of comparative data exists for languages across cultures, in terms both of structure and of use, no such similar body of data exists for music, which severely limits their comparability as communicative media. Although significant research has been undertaken that has elucidated many of the cognitive and neuroscientific underpinnings of human musicality, at present bodies of theory and batteries of tools that have broad applicability in identifying common characteristics across musical cultures, and that are widely accepted, do not yet exist. In the 1960s and 70, the American ethnomusicologist Alan Lomax embarked on the *Cantometrics* project (Lomax et al., 1978) which was intended to provide just such comparative data in respect of music; had this been successful, it would have provided secure bases for identifying commonalities and divergences in music across cultures, and for mapping the distribution of musical practices across geographical space and historical - and prehistorical - time. However, Lomax's methods and findings were severely criticised (see, e.g., Feld & Fox, 1994) and have never received broad acceptance. Nevertheless, new statistical techniques have recently been applied to Lomax's coded data which have already yielded suggestive hypotheses (Leroi & Swire, 2006) about the relationships between geographical distributions of cross-cultural musical types and historical and evolutionary processes. Having said this, several of the criticisms raised in respect of the original project remain. Were the bases for these criticisms successfully to be addressed, we would be able to trace the relationships between language and music, and the evolution of human systems of communication, with a much greater degree of certainty.

## References


Albersnagel, F. A. (1988). Velten and musical mood induction procedures: a comparison with accessibility of thought associations. *Behaviour Research and Theory, 26*(1), 79-96.

Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences, 28*(2), 105-124.

Auer, P., Couper-Kuhlen, E., & Müller, F. (1999). *Language in time: the rhythm and tempo of spoken interaction* New York: Oxford University Press.

Basso, E. (1985). *A musical view of the universe*. Philadelphia: University of Pennsylvania Press.

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

Blacking, J. (1976). *How musical is man?* London: Faber.

Blacking, J. (1995). *Music, Culture and Experience*. London: University of Chicago Press.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, 98*(20), 11818-11823.

Bolinger, Dwight (1978). Intonation across languages. In J.H. Greenberg (Ed.) *Universals of human language*, 471-524. Stanford: Stanford.

Boltz, M. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception, 18*, 427-454.

Boltz, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory & Cognition, 32*(7), 1194-1205.

Call, J., & Tomasello, M. (2005). What chimpanzees know about seeing revisited: an explanation of the third kind. In N. Eilan, C. Hoerl, T. McCormack & J. Roessler (Eds.), *Joint attention: communication and other minds*, 45-64. Oxford: Oxford University Press.

Clendinnen, I. (2005). *Dancing with strangers: Europeans and Australians at First Contact.* Cambridge: Cambridge University Press.

Cohen, D., & Inbar, E. (2001). Musical imagery as related to schemata of emotional expression in music and on the prosodic level of speech. In R. I. Godøy & H. Jørgenson (Eds.), *Musical Imagery*, 137-159. Lisse: Swets & Zeitlinger B.V.

Cohen, D. J. (2001). "The Imperfect Seeks Its Perfection": Harmonic Progression, Directed Motion, and Aristotelian Physics. *Music Theory Spectrum, 23*(2), 139-169.

Conard, N. J., & Bolus, M. (2003). Radiocarbon dating the appearance of modern humans and timing of cultural innovations in Europe: new results and new challenges. *Journal of Human Evolution, 44*, 331–371.

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

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

Couper-Kuhlen, E. (1993). *English speech rhythm: form and function in everyday verbal interaction* Amsterdam: J. Benjamins.

Cox, A. W. (2001). The mimetic hypothesis and embodied musical meaning. *5*, 195-209.

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

Cross, I. (2005). Music and meaning, ambiguity and evolution. In D. Miell, R. MacDonald & D. Hargreaves (Eds.), *Musical Communication*, 27-43. Oxford: Oxford University Press.

Cross, I. (2006). Music and social being. *Musicology Australia, 28*, 114-126.

D'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A.-M., Soressi, M., et al. (2003). Archaeological evidence for the emergence of language, symbolism, and music - an alternative multidisciplinary perspective. *Journal of World Prehistory, 17*(1), 1-70.

D'Errico, F., & Villa, P. (1997). Holes and grooves: the contribution of microscopy and taphonomy to the problem of art origins. *Journal of Human Evolution, 33*(1), 1-31.

Darwin, C. (1998). *The expression of the emotions in man and animals* (3rd ed.). London: HarperCollinsPublishers.

Davies, J. B. (1978). *The psychology of music*. London: Hutchinson.

Deliège, I., & Sloboda, J. A. (Eds.). (1996). *Musical beginnings*. Oxford: Oxford University Press.

Demany, L., & Clément, S. (1995). The perception of frequency peaks and troughs in wide frequency modulations. II. Effects of frequency register, stimulus, uncertainty, and intensity. *Journal of the Acoustical Society of America, 97*(4), 2454-2459.

Demany, L., & McAnally, K., I. (1994). The perception of frequency peaks and troughs in wide frequency modulations. *Journal of the Acoustical Society of America, 96*(2), 706-715.

Dilley, L. C., & McAuley, J. D. (2007: in revision). Distal prosodic effects on word segmentation and lexical processing. *Journal of Memory and Language*.

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

Dunbar, R. (1996). *Grooming, gossip and the evolution of language*. Cambridge, MA: Harvard University Press.

Evers, S., & Suhr, B. (2000). Changes of the neurotransmitter serotonin but not of hormones during short time music perception. *European Archives of Psychiatry and Clinical Neuroscience, 250*(3), 144-147.

Feld, S. (1982). *Sound and sentiment: birds, weeping, poetics and song in Kaluli expression.* Philadelphia: Publications of the American Folklore Society. New series; 5. BZVDW

Feld, S. (1996). Pygmy POP. A genealogy of schizophonic mimesis. *Yearbook for Traditional Music*(28), 1-35.

Feld, S., & Fox, A. A. (1994). Music and language. *Annual Review of Anthropology, 23*, 25-53.

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

Foley, R. A. (1995). *Humans before humanity.* Oxford: Blackwell.

Gopnik, A. (1999). Theory of mind. In R. A. Wilson & F. C. Keil (Eds.), *MIT encyclopedia of cognitive sciences*, 838-841. Cambridge, MA: MIT Press.

Gorbman, C. (1987). *Unheard Melodies*. Bloomington: Indiana University Press.

Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology, *Speech Prosody 2002: An International Conference*. Aix-en-Provence, France.

Gussenhoven, C. (2005). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it and how did it evolve? *Science*, 1569-1577.

Hockett, C. F. (1960). Logical considerations in the study of animal communication. In W. E. Lanyon & W. N. Tavolga (Eds.) *Animal Sounds and Communication*, 392–430. American Institute of Biological Sciences. Washington, D.C.

House, J. (2006). Constructing a context with intonation. *Journal of Pragmatics, 38*(10), 1542-1558.

Hughes, D. (1991). Grammars in non-Western musics: a selective survey. In P. Howell, R. West & I. Cross (Eds.), *Representing musical structure* (pp. 327-362). London: Academic Press.

Huron, D. (1996). The melodic arch in western folksongs. *Computing in Musicology, 10*, 3-23.

Huron, D. (2006). *Sweet anticipation: music and the psychology of expectation*. Cambridge, MA.: MIT Press.

Huron, D., Kinney, D., & Precoda, K. (2006). Influence of pitch height on the perception of submissiveness and threat in musical passages. *Empirical Musicology Review 1*(3), 170-177.
Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception, 20*(2), 151-171.
Jones, Mari Riess, and Boltz, Marilyn. 1989. Dynamic attending and responses to time. *Psychological Review* 96, 459-491.
Juslin, P. (2001). Communicating emotion in music performance. In P. Juslin & J. A. Sloboda (Eds.), *Music and emotion: theory and research*, 309-337. Oxford: Oxford University Press.
Juslin, P., & Sloboda, J. A. (Eds.). (2001). *Music & emotion: theory and research*. Oxford: OUP.
Juslin, P. N., & Laukka, P. (2003). Evidence of cross-modal similarities. *Annual New York Academy of Science, 1000*, 279-282.
Kassabian, A. (2001). *Hearing film: tracking identifications in contemporary Hollywood film music*. London: Routledge.
Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.
Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., and Friederici, A. D. 2005. Perception of phrase structure in music. *Human Brain Mapping* 24:259-273.
Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3), 396-414
Ladd, D. R. (1984). Declination: a review and some hypothesis. *Phonology Yearbook, 1*, 53-74.
Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
Leroi, A., & Swire, J. (2006). The recovery of the past. *The World of Music, 48*(3), 43-54.
Lomax, A., Rudd, R., Grauer, V. A., Berkowitz, N., Hawes, B. L., & Kulig, C. (1978). *Cantometrics: an approach to the anthropology of music: audiocassettes and a handbook*. Berkeley: University of California Extension Media Center.
MacDonald, R., Hargreaves, D., & Miell, D. (Eds.). (2002). *Musical Identities*. Oxford: Oxford University Press.
Malinowski, B. (1974). *Magic, science, and religion, and other essays*. London: Souvenir Press.
Marrett, A. (2005). *Songs, Dreamings, and Ghosts: the Wangga of North Australia*. Hanover, CT: Wesleyan University Press.
Miell, D., MacDonald, R., & Hargreaves, D. (Eds.). (2005). *Musical Communication*. Oxford: Oxford University Press.
Mithen, S. (2005). *The singing Neanderthals: the origins of music, language, mind and body*. London: Weidenfeld & Nicholson.
Montague, R. (1974) *Formal philosophy: selected papers*. New Haven: Yale University Press.
Morley, I. (2003). *The evolutionary origins and archaeology of music: an investigation into the prehistory of human musical capacities and behaviours.* Unpublished Ph.D, University of Cambridge, Cambridge.

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist, 111*(981), 855-869.
Nettl, B. (2005). *The study of ethnomusicology: thirty-one issues and concepts* (2nd ed.). Urbana & Chicago: University of Illinois Press.
Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica, 41*(1), 1-16.
Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols & J. J. Ohala (Eds.), *Sound Symbolism*, 325-347. Cambridge: Cambridge University Press.
Owings, D. H., & Morton, E. S. (1998). *Animal vocal communication: a new approach.* Cambridge: CUP.
Owren, M. J., & Rendall, D. (2001). Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology, 10*, 58–71.
Palmer, C., & Hutchins, S. (2006). What is musical prosody. *The Psychology of Learning and Motivation, 46*, 245-278.
Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 564-573.
Repp, B. H. (2006). Musical Synchronization. In E. Altenmuller, M. Wiesendanger & J. Kesserling (Eds.), *Music, Motor Control, and the Brain*, 55-76. Oxford: Oxford University Press.
Richards, P. (2007). The Emotions at War: A Musicological Approach to Understanding Atrocity in Sierra Leone. In Perri 6, C. Squire, S. Radstone and A. Treacher (Eds.) *Public Emotions*, (pp62-84). Basingstoke: Palgrave-Macmillan.
Richardson, D. C., Dale, R., & Kirkham, N. Z. (2006). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*(5), 407-413.
Rickard, N. S., Toukhsati, S. R., & Field, S. E. (2005). The effect of music on cognitive performance: insight from neurobiological and animal studies. *Behavioral and Cognitive Neuroscience Review, 4*(4), 235-261.
Russo, F. A., and Cuddy, L. L. . 1999. Motor theory of melodic expectancy. In *Acoustical society of America ASA/EAA/DAGA '99 meeting lay language papers*. Berlin.
Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*(3), 235-248.
Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology, 54*, 145-173.
Shennan, S. (2002). *Genes, memes and human history*. London: Thames and Hudson.
Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual Interpersonal Postural Constraints Are Involved in Cooperative Conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326–332.
Slobin, M. (1993). *Subcultural sounds: micromusics of the West.* Hanover: Wesleyan University Press.

Sloboda, J. A., O Neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the Experience Sampling Method. *Musicae Scientiae, 5*(1), 9-32.
Small, C. (1998). *Musicking*. London: Wesleyan University Press.
Sperber, D., & Wilson, D. (1995). *Relevance: communication and cognition*. Oxford: Blackwell Publishers Ltd.
Stevens, C. (2004). Cross-cultural studies of musical pitch and time. *Acoustics, Science and Technology, 25*(6), 433-438.
Számadó, S., & Szathmáry, E. (2006). Selective scenarios for the emergence of natural language. *Trends in Ecology and Evolution, 21*(10), 555-561.
Thomassen, J. M. (1982). Melodic accent: experiments and a tentative model. *Journal of the Acoustical Society of America, 71*(6), 1596-1605.
Thompson, W. F., & Balkwill, L.-L. (2006). Decoding speech prosody in five languages. *Semiotica, 158*(1/4), 407-424.
Titon, J. T. (Ed.). (1996). *Worlds of music: an introduction to the music of the world's peoples*. New York: Schirmer Books.
Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs: Prentice Hall.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675-691.
Trevarthen, C. (1999/2000). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientiae, special issue: Rhythm, musical narrative, and the origins of human communication*, 155-215.
Turnbull, C. M. (1965). *Wayward servants: the two worlds of the African pygmies*. New York: Natural History Press.
Watkins, A. J., & Dyson, M. C. (1985). On the perceptual organisation of tones sequences and melodies. In P. Howell, I. Cross & R. West (Eds.), *Musical Structure and Cognition*, 71-119. London: Academic Press.
Watt, R. J., & Ash, R. L. (1998). A psychological investigation of meaning in music. *Musicae Scientiae, 2*(1), 33-54.
Wennerstrom, A. (2001). *The music of everyday speech: prosody and discourse analysis*. Oxford: Oxford University Press.
Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics, 38*(10), 1559-1579.
Wilson, M., & Wilson, T. (2005). An oscillator model of turn-taking. *Psychonomic Bulletin and Review, 12*(6), 957-968.
Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication, 18*(46-67).
Yallop, C., and Clark, J. (1990). *An introduction to phonetics and phonology*. Oxford: Basil Blackwell.
Zhang, J., Harbottle, G., Wang, C., & Kong, Z. (1999). Oldest playable musical instruments found at Jiahu early Neolithic site in China. *Nature, 401*, 366–368.