

HUMAN NATURE REVIEW

ISSN 1476-1084
<http://human-nature.com/>

Book Review

Synch, Song, and Society

A review of *The Singing Neanderthals: The Origins of Music, Language, Mind and Body* by Steven Mithen, Weidenfeld & Nicholson, 2005; ISBN: 0297643177

by

William L. Benzon

There are at least two reasons why an intellectual specialist writes for a general audience, including intellectual specialists from other disciplines. One is to contribute to civic life by explaining difficult but important subjects in a way that makes them accessible to the citizen who is curious about the world. Steven Pinker's *The Language Instinct* is a distinguished recent example of such a book. But a specialist in some discipline may also seek a broader canvas than is available within the guiding principles of the specialized journal article or professional monograph. Here I think of Jared Diamond's *Guns, Germs and Steel*. When done well, a book of this type has a value for the specialist that the unadorned popularization, no matter how well done, does not have.

There are important problems that cannot be handled within confines of a single intellectual discipline. The origins of humankind is one of these problems. No matter which facet of that problem interests you, you inevitably find yourself looking at everything – or so it seems. Is music an offshoot of language or did a music-like activity evolve prior to language? In principle we could answer this question by traveling back in time and making direct observations. Unfortunately, that particular principle cannot be realized in the world as we know it, so we must instead approach human origins indirectly by gathering evidence from a wide variety of disciplines – archeology, physical and cultural anthropology, cognitive psychology, developmental psychology, and the neurosciences – and piecing it together. Such work entails a level of speculation that is incompatible with the publication demands of the specialist literature. It also demands a breadth of

knowledge that is all but impossible. When done well, however, such a book contributes to specialist investigations by establishing a framework within which more detailed work can be done.

Steven Mithen's current book, *The Singing Neanderthals*, is in this vein. He presents a wide range of evidence and ideas in clear and accessible prose that argues convincingly that the evolution of language was preceded by something that is neither language nor music as we now know them. Mithen has coined two neologisms for this, Hmmm (Holistic, multi-modal, manipulative, and musical) and Hmmmm (Holistic, multi-modal, manipulative, musical, and mimetic). While the full expansions of Mithen's neologisms give you a pretty good idea of what he is talking about, the neologisms themselves are too cutesy for my taste and simply do not work well as words. I prefer a different neologism, Christopher Small's (1998) *musicking*, from "to music." This neologism is not so cute and means music and/or dance or both and anything deeply connect with them. Words with roughly that meaning are common enough, but not in our culture. Björn Merker (2000, 320) has noted that the Greek *mousiké* encompasses melody, dance, and poetry, while the Bantu *ngoma* covers drumming, singing, dancing and festivity and the Blackfoot *saapup* covers singing, dancing, and ceremony. Where specificity is important and context alone is not sufficient to convey my meaning, I will supply appropriate qualifiers.

After an introductory chapter, Mithen has six chapters that are more or less about music and the brain. The second part of the book consists of ten chapters starting with a review of primate communication and then presenting Mithen's speculative reconstruction of the steps leading through the origins and elaboration of musicking to its final differentiation into language and music. All of these chapters contain useful information and intelligent synthesis that is well and cleanly presented to a cumulatively brilliant effect. Rather than attempt to review all of this material I will only cover the topics that, for better or worse, I found interesting. I will then offer some speculations of my own on music and sociality and conclude with some general remarks.

SELECTED TOPICS: AFFECT REGULATION, GROUP NORMS, AND LANGUAGE

Before picking up Mithen's argument in the sixth chapter, "Talking and singing to baby," I will briefly characterize the first five. The first chapter introduces the book and notes that music has been unfairly neglected by evolutionary scholars. The second chapter compares and contrasts music with language and introduces the idea of a common evolutionary precursor to both. The third and fourth chapters focus on the neural underpinnings and behavioral expression of music and language, respectively, considering the effects of brain

damage that diminishes one or the other, but not both.

Opening chapter five with the work of Isabelle Peretz (pp. 62 ff.), Mithen summarizes what is known about music and language in the brain, making the point that they share some circuits, while other circuits seem specialized to each. As music processing is widely distributed in the brain, it makes no sense to talk of a music module. Rather, if one is to talk of modules, one should talk of modules for specific functions, e.g. interval analysis, rhythm analysis, and vocal plan formation.

Note that, in the preceding paragraph, I cited Isabelle Peretz without indicating a specific source. I did that because Mithen discussed her work at some length. In the following discussions I cite with other researchers in same way.

Infants and Affect

In chapter six, “Talking and singing to baby,” Mithen examines mother-infant interaction, developing a theme from Colin Trevarthen, that mother and infant interact in an exquisitely timed multi-modal dance that is a necessary foundation for successful social interaction. Mithen begins with “infant-directed speech” (IDS), also-known-as baby talk or motherese, in which the musical elements of speech are exaggerated: wider pitch range, longer vowels, exaggerated pauses, and more repetition. Adults and children automatically adopt IDS when interacting with infants, who much prefer it to normal speech. He cites the work of Anne Fernald who has found that IDS progresses through four stages. In the first three stages IDS helps the mother to maintain contact with the infant, to modulate arousal and emotion, and to convey feelings and intentions to the infant. It is only in the fourth stage that IDS is explicitly directed at facilitating language learning. IDS appears to be universal, with some variations to accommodate languages, such as Chinese, where pitch is phonemically meaningful.

Singing to infants is universal as well, and is even more emotionally potent, a fact that Mithen will evoke later on when he offers Ellen Dissanayake’s argument that singing evolved to enhance mother-infant interactions as infants were born earlier in the life cycle. Consequently they required even more care and that care was required for a longer period of time. Music strengthened the relationship between mother and child and contributed to the child’s enculturation. Singing would also have allowed mothers to attend to infants who had been put down while mother attended to other tasks – an hypothesis argued by Dean Falk.

These discussions lead naturally to chapter seven, on music and emotion, where Mithen reviews the position advocated by Keith Oatley and Philip Johnson-Laird, “that emotions guide action in situations of imperfect knowledge and multiple, conflicting goals” (87), a position, incidentally, that bears

comparison with Warren McCulloch's model of the reticular activating system (Kilmer, McCulloch, and Blum 1969). Since "the social world provides the greatest cognitive challenge to humans beings" it follows that "our more complex emotions relate directly to our social relationships" (87). Mithen presents experimental evidence that music accurately conveys emotions such as happiness, sadness, and anger. In experimental settings people with no particular musical training judged such matters as reliably as trained musicians, and women did slightly better than men (93). Mithen then reviews work on music therapy, experiments by Alice Isen showing that "happiness enables people to think more creatively" (98) and a study by Rona Fried and Leonard Berkowitz showing that people who had listened to soothing music were more helpful than controls and than people who had listened to exciting music or sad music (99-100).

Walking and Rhythm

To my mind, perhaps the most critical issue in understanding music and its evolution is that of rhythm, for rhythm sets the terms through which the actions of individuals become mutually coupled when they are musicking (Benzon 2001, 23-68, 116-142). Mithen argues that our rhythmic capacities fell into place roughly 1.8 million years ago with the emergence of full bipedalism in *Homo ergaster*. Following Leslie Aiello he points out that "Standing or walking on two legs requires that the center of gravity is constantly monitored and small groups of muscles frequently recruited and changed to correct its position; the movement of the legs has to be integrated with that of the arms, hands, and trunk in order to maintain a dynamic balance" (146). This would have required greater rhythmic coordination among muscle groups and hence a larger brain to manage that coordination.

In support of this emphasis, Mithen cites the work of Michael Thaut, who has been investigating the value of rhythmic sound in treating patients with Parkinson's disease, which typically disturbs motor control. Thaut discovered that patients who received "gait-training" sessions that involved rhythmic auditory cues did better than those whose training did not involve auditory cueing; they, in turn, did no better than those who received no training at all. Somehow – the mechanism is not known – rhythmic sound is able to substitute for endogenous rhythm lost through disease. (The results of the training, however, disappeared after five weeks with no training at all.)

Mimetic Culture and Holistic Utterance

In chapter eleven, "Imitating nature," Mithen introduces Merlin Donald's (1991) influential concept of mimetic culture as an intermediate form between the

episodic culture of apes and the mythic culture of modern humans. As the name suggests, mimetic culture is based on the ability to imitate and allowed *Homo ergaster*, *Homo erectus*, and *Homo heidelbergensis* “to create new types of tools, colonize new landscapes, use fire and engage in big-game hunting” (Mithen 167). Mimetic humans would have been able to mime the vocalizations of animals and other humans for their play and amusement, for ritual, and for practical communication as well.

These early humans would have had a large repertoire of holistic utterances, each functioning as a complete message – a point argued by Alison Wray. They would not have been able to segment these messages into components and then recombine the components into new utterances with new meanings – as we do with words – but the individual messages would have considerably larger and richer than that of the call systems of apes. (Note that Mithen has thus explicitly rejected the arguments of Derek Bickerton, who imagines proto-language to have been combinatorial in nature, but operating from a restricted range of words and syntax.) Because each holistic phrase would have been long and complex, it would have been difficult to introduce new phrases. Hence this system would have “been dominated by utterances descriptive of frequent and quite general events” and have been quite conservative.

Obviously we have no record of these utterances, but the archeological record does have indications of cultural conservatism. The repertoire of stone tools was both limited and unchanged between 1.8 and 0.25 million years ago; Mithen gives particular emphasis to the constant form of hand-axes (164). Mithen suggests that, because their finely wrought form exceeds the practical demands of butchery, wood-working, and cutting plants, these hand-axes may have been fitness indicators in the sort of sexual selection regime Geoffrey Miller has advocated.

Beyond this, I note that Ralph Holloway (1969, 1981) long ago suggested that strongly-conserved hand-axe form was an indicator of social norms. Those forms could not be conserved from one generation to the next unless there was a deliberate intention to do so. One has to note the significant features of an existing axe and discipline one’s knapping motions to produce that result. That is considerably more exacting than simply producing an axe with a sharp edge and appropriate heft. The motivation behind such exacting form, then, is not practical. Nor can it be merely aesthetic, which would allow for considerable individual variation. That leaves us with a desire to conform to social norms. Given the importance of such norms, that may in itself be a sufficient motivation for their form, to serve as a visible token of social solidarity. In any event, Holloway’s observation does not contradict Miller’s, and now Mithen’s hypothesis. Norms are norms, regardless of their specific purpose and norms that serve multiple ends are likely to be particularly strong.

Groups and Music

Mithen begins developing his argument about music in groups with a variation on Miller's argument about sexual selection. Perhaps, argues Björn Merker, males in one group would band together in choral musicking to attract lone females wandering away from their groups. He rejects the idea, however, on the grounds that such musicking would also attract predators. He does suggest, however, that such activity might well build trust among the males, and females as well. From there he discusses the ideas of historian William McNeill (1995), who emphasizes the value of dance and drill in creating cohesion within the group and talks as well how group musicking leads to a blurring of personal boundaries and a consequent sense of merging with the group.

Mithen then takes up a discussion of Robert Axelrod's classic work on the prisoner's dilemma, which shows the value of cooperation, but also the difficulties of achieving it in a world where people's interests are neither completely consonant nor opposed and where it is difficult to know one another's intentions. Mithen (214) suggests that

music-making is a cheap and easy form of interaction than can demonstrate a willingness to cooperate and hence may promote future cooperation when there are substantial gains to be made ... It can be thought of as the first move of a 'TIT for TAT' strategy that is to always cooperate, one that can be undertaken at no risk because there is nothing to lose if the other members defect – that is, if they do not join in the song or dance.

The problem with this, Mithen notes, is that, because the costs are low and the benefits are high, such musicking could easily be exploited by free-riders. To the extent, however, that musicking leads to a merging of self with the group – McNeill's argument – this problem may be dissolved (215):

Those who make music together will mould their own minds and bodies into a shared emotional state, and with that will come a loss of self-identity and a concomitant increase in the ability to cooperate with others. In fact, 'cooperate' is not quite correct, because as identities are merged there is no 'other' with whom to cooperate, just one group making decisions about how to behave.

Thus musicking leads to the formation of a group identity within the scope of which individuals will be strongly inclined toward cooperation to their mutual benefit.

Neanderthals

With this argument behind him, Mithen takes up the case of the singing Neanderthals (chapter 15), reprising his 1996 argument that, while quite intelligent, their intelligence was domain-specific. They could not make connections between knowledge of the natural world, physical materials, and social interaction (232). He feels, however, that this is not adequate to account for their large brains. Just why Mithen believes this is not at all clear since he has not made a quantitative argument about the relationship between brain size and intelligence. Nonetheless he goes on to argue that, while Neanderthals did not have language, they would most likely have had a rich repertoire of holistic utterances of the sort advocated by Alison Wray (234):

The Neanderthals would have had a larger number of holistic phrases than previous species of *Homo*, phrases with greater semantic complexity for use in a wider range of more specific situations. It think it is also likely that some of these were used in conjunction with each other to create simple narratives.

Noting that life was difficult for the Neanderthals, with few living beyond the age of 35, Mithen (236) suggests that there “is unlikely ever to have been a population of humans – modern, Neanderthal or otherwise – for whom the creation of social identity to override that of the individual was more important. For that, music is likely to have been essential.” This is certainly the case for modern humans. When times are tough, we make music, which enables social bonding and thus facilitates mutual support.

Mithen then takes up, in order, discussions of mimesis and hunting, stone tools and mating, infants, burial, performance spaces, and a flute-like artifact discovered in Slovenia, which he suggests is not a flute. He concludes by suggesting that, in comparison to us, the Neanderthals were linguistically challenged but musically advanced, an argument he bolsters by an assertion and two anecdotes. The assertion is that “the evolution of language has inhibited the musical abilities inherited from the common ancestor we share with *Homo neanderthalensis*” (245). As for the anecdotes, one is about a particularly radiant ballet performance he had attended and another about walking with a musical savant he had discussed early in the book.

Mithen’s belief that language evolution inhibits music seems to be based on evidence he presented earlier (76-79) that we are born with absolute pitch, though the vast majority of us lose it subsequent to language acquisition. But Mithen presents no evidence that absolute pitch is important in musical ability. Though musicians are more likely to have absolute pitch, most lack it. Nor, so far

as I can tell, is there good reason to think it an important component of musical skill (Sloboda 1985, 176-178). As for Mithen's anecdotes, they are just that, anecdotes; and they have an air of romanticism about them, about how certain exotic Others – Neanderthals, savants – are more deeply in touch with the sensory world than “we” are.

Still, my skepticism about this matter – which would seem to be the justification for the book's title – does not arise from any specific reason to doubt that Neanderthals had advanced musical abilities. I simply do not know. I keep wondering what “our” musical abilities would be if we had to make our own music instead of being able so easily to listen to recordings and broadcasts. For me that question simply “washes out” Mithen's speculations about Neanderthal musical superiority.

But these are relatively minor quibbles about what is, in the full scope of the book, a relatively minor point. Whether or not the Neanderthals were musically superior to *Homo sapiens*, Mithen's arguments are important.

Language

In his penultimate chapter (16) Mithen discusses the final emergence of biologically modern language, and with it, cognitive fluidity – the ability to integrate ideas and information across the physical material, natural, and social domains. The final chapter (17) covers the dispersal of humankind around the globe, the remains of holistic speech within music and segmented language, and reviews the cultural uses of music.

These are interesting chapters and I recommend them; but I see little need to review these arguments. Most of Mithen's work has been done by the end of the Neanderthal chapter. What Mithen has established by that point is that Early Humans – Neanderthals as well as others – would have developed a relatively rich and complex culture *prior to the final emergence* of biologically modern man and culture. Just how rich is not, of course, certain. Mithen denies that Neanderthals were capable of symbolic behavior; I am not so sure and will shortly suggest that group musicking would have led to symbolic behavior in groups that did it – including, of course, the Neanderthals.

It is clear, however, that Mithen's argument requires the development of segmented language in the final stage of the evolutionary trajectory leading to *Homo sapiens*. That final evolutionary step may also require the development of syntax and of semantics depends on one's theory of language. Exactly what is required is at issue in the debate initiated by Hauser, Chomsky, and Fitch (2002) in which a distinction is made between a broadly constructed language faculty and a narrowly construed one. Hauser, Chomsky, and Fitch argue that recursion is the only feature of the narrowly construed faculty and is most likely the only one

unique to humankind.

Given that late pre-humans had a rich repertoire of holistic utterances, I suspect that segmentation may have been the major requirement and that syntax and semantics would have followed from that segmentation (cf. Benzon and Hays 1988). I adhere to a view – often associated with cognitive linguistics – that regards semantics as being grounded in a highly structured repertoire of cognitive and sensorimotor schemas. Given segmentation, and the association of specific segments with specific perceptual or cognitive schemas, the job of syntax is to arrange segments in sequential order. In this view syntactic structure is, in fact, driven by cognitive or semantic structure.

Whatever the case may be, this is not the place to engage in that debate, which is not about music. Mithen has gathered an impressive array of evidence, phylogenetic, ontogenetic, behavioral, and anatomical, both body and brain. While I find his argument impressive, I am strongly biased in his favor. Those who give priority to language, and argue that music is one of its by products, may not be convinced. Mithen's argument is built on speculative interpretations of bones, artifacts, and site topographies each of which can be replaced by alternative interpretations. None of these speculations, Mithen's or alternatives, is an adequate substitute for direct observation of behavior, which is not available to us.

This does not necessarily mean that we will never be able to resolve the question: Which came first, music or language? More evidence will be found and some of it may well favor one line of speculation over others. I think, however, we need a different kind of argument, one that is more deeply grounded in the intrinsic properties of the nervous system.

This kind of argument is about the evolutionary trajectory leading from the episodic intelligence that Merlin Donald (1991) has attributed to apes to the mythic intelligence he attributes to biologically modern humans. Is the nature of the nervous system such that that evolutionary trajectory *must* go through musicking on the way to language?

It is by no means clear to me that we are ready to address this question in any detail. But I want to outline such an argument in the next two sections. First, following Walter Freeman, I will argue that the critical factor is getting two or more individuals into the same *intentional framework*; without that, one cannot learn language from the another or use language to communicate. Then I will sketch out an evolutionary trajectory, building on Mithen's work and on my own work (Benzon 2001), and show how musicking can give rise to awareness of the group *as such* and to simple symbols. The argument needs to be elaborated in greater technical detail than I can provide here. But if I am correct in this outline, then the evolutionary trajectory from ape behavior to human behavior must pass through musicking before it arrives at biologically modern language.

SYNCHRONY AND INTENTIONALITY

In this discussion I will assume that the nervous system operates as a self-organizing dynamical system as, for example, Walter Freeman (1995, 1999, 2000b) has argued. Using Freeman's work as a starting point, I have previously argued that, when individuals are musicking with one another, their nervous systems are physically coupled with one another for the duration of that musicking (Benzon 2001, 47-68). There is no need for any symbolic processing to interpret what one hears or so that one can generate a response that is tightly entrained to the actions of one's fellows.

My earlier arguments were developed using the concept of coupled oscillators. The phenomenon was first reported by the Dutch physicist Christian Huygens in the seventeenth century (Klarreich 2002). He noticed that pairs of pendulum clocks mounted to the same wall would, over time, become synchronized as they influenced one another through vibrations in the wall on which they were. In this case we have a purely physical system in which the coupling is *direct* and completely mechanical.

In this century the concept of coupled oscillation was applied to the phenomenon of synchronized blinking by fireflies (Strogatz and Steward 1993). Fireflies are, of course, living systems. Here we have energy transduction on input (detecting other blinks) and output (generating blinks) and some amplification in between. In this case we can say that the coupling is *mediated* by some process that operates on the input to generate output. In the human case both the transduction and amplification steps are considerably more complex. Coupling between humans is certainly mediated. In fact, I will go so far as to say that it is mediated in a particular way: each individual is comparing their perceptions of their own output with their perceptions of the output of others. Let us call this *intentional* synchrony.

Further, this is a completely voluntary activity (cf. Merker 2000, 319-319). Individuals give up considerable freedom of activity when they agree to synchronize with others. Such tightly synchronized activity, I argued (Benzon 2001), is a critical defining characteristic of human musicking. What musicking does is bring all participants into a temporal framework where the physical actions – whether dance or vocalization – of different individuals are synchronized on the *same time scale* as that of neural impulses, that of milliseconds. Within that shared intentional framework the group can develop and refine its culture. Everyone cooperates to create sounds and movements they hold in common.

There is no reason whatever to believe that one day fireflies will develop language. But we know that human beings have already done so. I believe that, given the way nervous systems operate, musicking is a necessary precursor to the

development of language. A variety of evidence and reasoning suggests that talking individuals must be within the same intentional framework.

Consider an observation that Mithen offers early in his book (p. 17). He cites work by Peter Auer who, along with his colleagues, has analyzed the temporal structure of conversation. They discovered that, when a conversation starts, the first speaker establishes a rhythm to which the other speakers time their turn-taking. That is, even though they are only listening, other parties are actively attuned to the rhythm of the speaker's utterance (cf. Condon 1986). What if this were necessary to conversation, and not just an incidental feature of it?

Let us recall some passages from Eric Lenneberg's landmark review and synthesis, *The Biological Foundations of Language* (1967). While he does not address the issue of conversational turn-taking, he does devote the better part of chapter three to timing issues. He was particularly interested in problems arising from the fact that neural impulses travel relatively slowly and that the recurrent nerve, innervating the larynx, is over three times as long as the trigeminal branch innervating the one of the jaw muscles. It also has a smaller diameter, which means that impulses travel more slowly in it than in the trigeminal. The upshot, observes Lenneberg, is that "innervation time for intrinsic laryngeal muscles may easily be up to 30 msec longer than innervation time for muscles in and around the oral cavity." He goes on to observe: "Considering now that some articulatory events may last as short a period as 20 msec, it becomes a reasonable assumption that the firing order in the brain stem may at times be different from the order of events at the periphery" (96). It is on the basis of such considerations, which he discusses in some detail, that Lenneberg concludes: "rhythm is ... the timing mechanism which should make the ordering phenomenon physically possible" (119).

It follows from this that, if you wish your utterances to smoothly intercalate with those of others, you need to share their rhythms; that is the only way your conversational entrances will be appropriately timed. Still, this might merely be a conversational convenience, not a necessity. So, let us consider the problem of speech perception.

We know that, while we tend to hear speech as a string of discrete sounds, that is something of an illusion. Sonograms do not show the segmentation that we hear so easily (Lenneberg 93-94). The brain is doing some sophisticated analysis of the sound stream. Though I am not aware that anyone has investigated this, I can imagine that it would be very useful if the listener operated within the same temporal framework as the speaker. This might help with the segmentation. If this is so, rhythmic synchronization is no longer simply a feature of how the nervous system happens to operate. It becomes essential to being able to treat the speech stream as a string of phonemes; it is necessary to linguistic communication.

Let us push the argument a step further. For the last decade or so there has been considerable interest in the notion that people acquire a so-called theory of mind (TOM) early in maturation and that this TOM is critical to interpersonal interaction (see e.g. Baron-Cohen 1995). Gaze following is one behavior implicated in TOM. Humans beyond a relatively early age will follow the direction of one another's gaze. I would like to suggest that we notice gaze direction in people with whom we synchronize, but not otherwise.

Think about the perceptual requirements of noticing and tracking gaze direction. Even at conversational distance, another person's eyes are small in relation to the whole visual scene; thus the visual cues for gaze direction will also be small. Further, people in conversation are likely to be in constant relative motion with respect to one another. The motions may not be large – head turns and gestures, trunk motion – but they will be compounded by the fact that one's eyes are in constant saccadic motion. Synchronization would eliminate one component of relative motion between people and therefore simplify the process of picking up the minute cues signalling gaze direction. But if one cannot properly synchronize with others, then those cues will be more difficult to notice and track. Thus the capacity for interpersonal synchrony may be a prerequisite for the proper functioning of TOM circuitry.

In this light let us now consider Paul Bloom's (2000) recent work on language acquisition. He has demonstrated that young children do more than merely associate the words they hear with the objects and events to which they refer. Such associations are not sufficient. Rather children make inferences about speaker's intentions when listening to them and learning the meanings of words they use. In the current parlance, children use a so-called theory of mind (TOM) to infer what, of many immediate possibilities, the speaker's words refer to. Inferring another's intentions also plays a large role in Quine's (1960, 26 ff.) classic argument about radical translation.

The current evidence is that monkeys lack this TOM, while the behavior in which chimpanzees seem to deceive their fellows suggests that they may have it in some measure (Mithen 117-118). Sophisticated discussions of TOM (e.g. Baron-Cohen 1995) talk of component skills and separate, though linked, neural circuits. Until the specialists have this sorted out, I will proceed on the assumption that whatever it is that allows chimpanzees to deceive one another, it is not of a kind that supports intuition about the intention behind another's utterance.

In any event, Bloom's discussion is about ontogeny, not phylogeny. The phylogenetic case is even more behaviorally demanding. Human children acquire language in a world where they are surrounded by proficient speakers who take pains to make themselves understood and work hard to understand the child. When language first originated there would have been no mature speakers at all. But, if we follow Mithen, those speakers would have been proficient in a proto-

language having a rich repertoire of holistic phrases.

By that time, however, Early Humans would also have been walking upright and thus would have had – again following Mithen – the precise rhythmic control necessary for tight interpersonal synchronization. And that – tight synchronization through music – Freeman argues, is what puts two humans in the same intentional framework. As Freeman (1995, 131) says music:

is wordless, illogical, deeply emotional, and selfless in its actualization of transient and then lasting harmony between intentional structures. It works in the action-reafferece-perception cycle that provides for all human understanding, and it constructs the sense of trust and predictability in each member of the community on which social interactions are based.

Freeman suggests that “the techniques for working matter into useful forms must have required the prior existence of channels for communication to support the social interactions. These channels are intentional and not logical in nature.” Thus the emergence of biologically modern language must have been preceded by musicking, for language depends on the capacity for interpersonal synchrony that musicking engenders.

GROUPS, SYMBOLS, AND NORMS

Let us now take an informal look at how musicking can give rise to the ability to symbolize the group as such. In the following paragraphs I will use four simple diagrams to present the argument.

Consider Figure 1, which represents the mind of a single individual, *A*. The nodes *c* and *b* are the neural representations of other individuals whom *A* knows while the ego node (epsilon ϵ) represents *A*'s neural self (Damasio 1994).

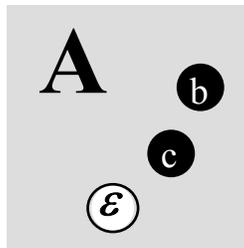


Figure 1: Mind of Individual A

Now examine Figure 2, representing three individuals interacting with one another in any fashion. We have *A* in the upper left and then *B* and *C* as well.

Notice that both B and C have neural representations of the other two, just like A, and that they too have neural selves.

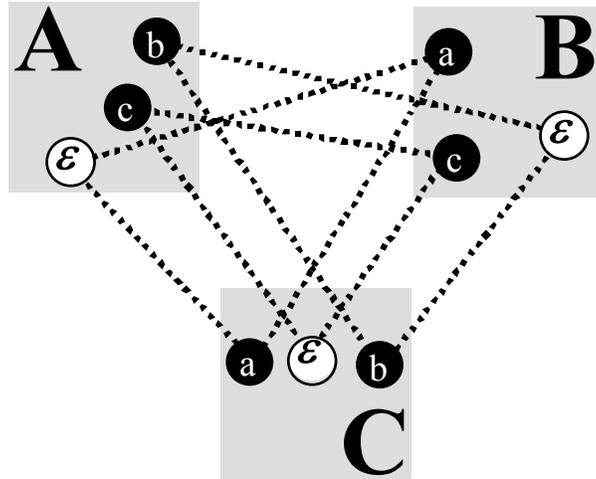


Figure 2: Three Individuals Interacting

Notice that I have used dotted lines to indicate corresponding neural representations. Thus A's epsilon node is associated with B's neural representation of A (that is, a) and to C's neural representation of A (that is, a). And so for B and C as well. Interactions among these individuals must necessarily be mediated by their sense of themselves and their sense of one another.

Let us imagine these individuals musicking together. C's own activity (dancing, gesturing, vocalizing, rattle-shaking, and so forth), for example, would be associated with C's epsilon node while the activities of the A and B would be associated with their representations in C. These would, in fact, be highly complex and distributed networks of neural activity. Each individual would have audio and visual activity representing what the other two are doing, but also his or her own activity (you can hear yourself and see your limbs and torso). Additionally, there would be motor and kinesthetic activity for only oneself – you cannot move another's muscles, nor experience movement in their joints. But if the group is synched up in a nice groove, then all of this neural activity, in all the sensory and motor channels, will have the same rhythmic pattern in common. If that rhythm came to be perceived as an entity in itself, that would be the grain of cognitive irritant around which individuals could construct the pearl of group awareness.

Consider the Figure 3, representing a group engaged in well-synchronized musicking:

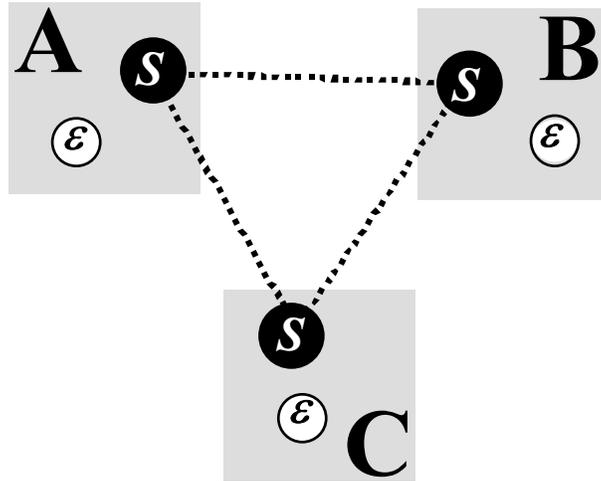


Figure 3: Synchronized Group

Here we have individuals A, B, and C. Each has neural self (ϵ), and they also have representations of one another (not depicted). The *S* nodes represent audio-motor networks used to create synchronized sound (by, e.g. clapping, foot stamping, whatever).

Now those *S* nodes are distinct neural entities within each individual. They are not identical to ego, nor to representations of one another. To be sure, they do not correspond to some object or symbol one can see and touch (e.g. a totemic statue); nor do they correspond to a name that can be uttered. Rather, they correspond to an activity that one can do, and do collectively. In a sense, the *S* nodes represent the collective activity of the group. They are the beginnings of individuals being able to conceptualize the group *as such*, which is more sophisticated than simply recognizing other individuals as group members and, on that account, according them appropriate treatment.

Now consider Figure 4:

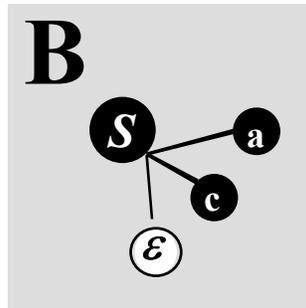


Figure 4: Awareness of the group as such

Figure 4 depicts B's mind as it represents the relationship between A, C, and ego (that is, B), on the one hand, and the act of synchronized musicking on the other. This need not represent B's mind during musicking – indeed, if B is deeply absorbed, he or she might not be aware of anything but the music that is the joint product of everyone's activity – but rather could represent B's recollection of or anticipation of musicking. Walter Freeman (2000a) has speculated that intense musicking induces Hebbian learning; perhaps that S node is the result of such learning.

Freeman's hypothesis, which Mithen discusses (216-217) is more specific than mere learning. It is about bonding. Freeman (2000a, 420) says:

What is at issue is the extent to which feelings of bonding and formation of a neural basis for social cooperation might be engendered by the same neurochemical mechanisms that evolved to support sexual reproduction in altricial species like ourselves, and that might mediate religious, political, and social conversions, involving commitment of the self to a person as in transference, fraternity, military group, sports team, corporation, nation, or new deity. The common feature is formation of allegiance and trust.

Freeman does not clearly distinguish between the mutual bonding of participants one to the other and the bonding of participants to some entity that somehow symbolizes the group. These are not, of course, mutually exclusive and, in fact, one can imagine that a process that starts in mutual bonding will, over time – which may be measured in weeks, years, or generations – result in bonding to a more abstract entity. This is what that S node is. That is to say, the biological mechanisms which bond infant to parent are now bonding individuals to symbolic representation of the group, a possibility that has been recognized in the attachment literature (Marris 1982).

It is not clear just how such a symbol would arise, how S becomes reified and conceptualized. Mithen discusses a site in Bilzingsleben in southern Germany dating roughly 400,000 years ago, for example, that is suggestive on this point. The site has “a vast quantity of animal bones, plant remains and stone artefacts, along with a few fragments of Early Humans” (178) and three circular arrangements of bones. Following Clive Gamble, Mithen suggests that setting up the anvil – a large stone or piece of wood used in manufacturing bone implements – was a major social act and that the circular sites were work and performance spaces. Proto-humans gathered around an anvil to make tools, butcher meat, and to eat. This activity would involve rhythmic striking and pounding and also, Mithen suggests, quasi-musical vocalizations, gestures, and dance moves. To the

extent that all this activity centers around an anvil, it might become the material “anchor” for S, the neural reflex of the synchronized group activity taking place around it. In discussing the Neanderthals he mentions a site at Bruniquel in southern France that may served a similar ceremonial purpose (242). At this point we are on a cultural evolutionary track that leads to religion (cf. Benzon 2001, 195-199).

Note that we are now considering dynamics on two levels; 1) that of individual nervous systems coupled in musicking, and 2) that of the entire life of the group. In the first case I am interested only in what happens between the beginning and conclusion of a session of musicking. In the second case I am thinking of the ongoing life of the group, both musicking and otherwise. From this point of view what we see is the group gathering together to engage in musicking and then dispersing as they go about their daily lives.

Considering the group as a dynamic system, these intervals of collective musicking are *attractors* in the life of the group (Benzon 2001, 154, 192-194). On these occasions the group members affirm their loyalty to the group; they constitute the group’s psychological home base. As such this larger self-organizing dynamic is at the heart of the group’s culture. This is the dynamic through which they together learn the central symbols and values of that culture. It is this larger cycle, for example, that affirms the importance and value of producing hand-axes to match the cultural standard, as Holloway has argued.

NEW ASSUMPTIONS ABOUT HUMAN NATURE

It is time to return where we began: the general audience. Mithen’s book speaks to the assumptions behind a long-entrenched intellectual culture that has privileged intellection over artistic expression and the individual over the group. While there has always been romantic opposition to the web of ideas woven on those assumptions, that opposition has recently begun to base its arguments on evidence from the neurosciences (e.g. Damasio 1996) and behavioral economics (e.g. Gintis and Bowles, et. al. 2005). That opposition need no longer trade on romanticism to advance its arguments. While the conventional wisdom is still dominant, it is beginning to look more and more like a deep-seated cultural bias rather than like conclusions to a well-reasoned argument based on solid empirical evidence.

When Steven Pinker wrote the final chapter of *How the Mind Works*, he assumed that conventional wisdom and argued that the arts are mental cheesecake (525): sweet, filling, flavorful, but of no adaptive value, however much cultural prestige they may have. That metaphor – mental cheesecake – set off a minor intellectual firestorm among humanists, especially musicologists and literary critics. Mithen obviously disagrees with Pinker with respect to music – the title of

his second chapter, “More than cheesecake?” takes aim at Pinker – and Mithen’s is by far the stronger argument. Joseph Carroll (1998, 2002) has made the argument on behalf of literature’s adaptive value. Only time will tell how this argument is going to fare.

But it may well be that more than purely conceptual matters are at issue. While I was researching my book on music I decided to read Russell A. Barkley’s review and synthesis of the ADHD literature, *ADHD and the Nature of Self-Control* (1997). Self-control was one of the topics that interested me both generally and in connection with music, and the book had been well-reviewed in *Science*. Barkley argued that ADHD does not involve inattention as much it involves poor self-control, which is a failure of some central executive function. In turn, Barkley asserts that the

. . . nature of this central executive . . . is time. More specifically, it is the conjecturing of the future that arises out of reconstruction of the past and the goal-directed behaviors that are predicated on these activities. Such activities . . . permit self-regulation relative to time. (p, 202)

Barkley goes on to point out that “time is an integral, inseparable part of the physical world” (p. 204), that “our will, therefore is . . . at time’s beck and call” (p. 205) and thus that “time, timing, and timeliness . . . become important concepts in understanding . . . goal-directed behavior and in determining it” (p. 209).

Yet nowhere does Barkley talk about music nor about music therapy. That is not particularly surprising, as the concern about ADHD is mostly about the ability of children to absorb the intellectual material taught in school. That, so the conventional wisdom has it, bears no relationship to music. Music is about feeling, about play, about fun, not about reasoning and intellect. The conventional wisdom would not lead one to think about music in connection with research directed toward improving school performance.

If Mithen and Trevarthen and McNeil and Dissanayake and a growing list of others are correct, however, then dance and music are deeply rooted in human evolutionary history. They belong to our biological nature. And we are now a century into a world where live musicking has largely been replaced by recorded and broadcast sound. What if the lack of sufficient active musicking at an early age is a causal factor in ADHD? If Barkley is correct, that timing is the problem in ADHD, then that possibility should be investigated. If the brain does NOT get sufficient musicking at a young age, then perhaps further maturation unfolds with timing difficulties built-in to its microstructure as a result of synaptic pruning in that music-poor environment.

I do not know whether or not this is true. What bothers me, however, is that, so far as I know, the matter is not being investigated. Conventional wisdom stands in the way of even considering the hypothesis, much less seriously investigating it.

This particular component of conventional wisdom may be costly. A recent NIH consensus panel estimated that ADHD cost public schools in the United States over \$3 billion in 1995, noting that “ADHD, often in conjunction with coexisting conduct disorders, contributes to societal problems such as violent crime and teenage pregnancy” (National Institutes of Mental Health 1998, p. 9). The cost of ADHD is high and we do not know, in the end, whether our treatment regimes are more efficacious than not. How much would it cost to investigate possible relationships between music, music therapy, and ADHD?

The case for musicking, and for the arts as well, has even more general implications. It goes to the heart of how we see ourselves, our society, and our place in the natural world. The argument is important for all of us.

The conventional view places a gulf between humankind and the natural world. We are rational, it is unthinking, and society is but a precariously ordered collection of selfish breeders. What if that is not so? What if society is the natural life-form for clever apes who learned to sing and dance before they learned to analyze and rationalize? Is it time for us to reconsider our nature, our relationships with one another, and our place in the world?

Author

William L. Benzon, 708 Jersey Avenue, 2A, Jersey City, NJ 07302, USA.

Acknowledgments

I would like to thank Walter Freeman, Ralph Holloway, Charlie Keil, and Tim Perper for their discussion of these issues. I take responsibility for remaining errors.

References

- Barkley, R. A. (1997). *ADHD and the Nature of Self Control*. New York, The Guilford Press.
- Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA, MIT Press.
- Benzon, W. L. and D. G. Hays (1988). "Principles and Development of Natural Intelligence." *Journal of Social and Biological Structures* **11**: 293-322.
- Benzon, W. L. (2001). *Beethoven's Anvil: Music in Mind and Culture*. New York, Basic Books.

- Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MIT Press.
- Carroll, J. (1998). "Steven Pinker's Cheesecake for the Mind." *Philosophy and Literature* 22: 478-485.
- Carroll, J. (2002). "Adaptationist Literary Study: An Emerging Research Program." *Style* 36: 596-617.
- Condon, W. S. (1986). Communication: Rhythm and Structure. *Rhythm in Psychological, Linguistic and Musical Processes*. J. R. Evans and M. Clynes. Springfield, Illinois, Charles C Thomas • Publisher: 55-78.
- Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, Avon Books.
- Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA, Harvard University Press.
- Freeman, W. J. (1995). *Societies of Brains: A Study in the Neuroscience of Love and Hate*. Hillsdale, NJ, Lawrence Erlbaum.
- Freeman, W. J. (1999). *How Brains Make Up Their Minds*. London, Weidenfeld and Nicholson.
- Freeman, W. J. (2000a). A Neurobiological Role of Music in Social Bonding. *The Origins of Music*. N. L. Wallin, B. Merker and S. Brown. Cambridge, MA, MIT Press: 411-424.
- Freeman, W. J. (2000b). *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics*. London, Springer-Verlag.
- Gintis, H., S. Bowles, et al. (2005). *Moral Sentiments and Material Interests: On the Foundations of Cooperation in Economic Life*. Cambridge, MIT Press.
- Hauser, M. D., N. Chomsky and W. T. Fitch (2002). "The Faculty of Language: What is It, Who Has it, and How Did It Evolve?" *Science* 298(5598): 1569-1579.
- Halloway, R. L. (1981). Culture, Symbols, and Human Brain Evolution: A Synthesis. *Dialectical Anthropology* 5: 287-303.
- Halloway, R. L. (1969). Culture: A Human Domain. *Current Anthropology* 10(4): 395-407.
- Kilmer, W. L., W. S. McCulloch and J. Blum (1969). "A Model of the Vertebrate Central Command System." *International Journal of Man-Machine Studies* 1: 279-309.
- Klarreich, E. (2002). "Huygens's Clocks Revisited." *American Scientist* 90(4). <http://www.americanscientist.org/Issues/Sciobs02/02-07sciobslocks.html>
- Lenneberg, E. (1967). *The Biological Foundations of Language*. New York, John Wiley & Sons. Inc.
- Marris, P. (1982). Attachment and Society. *The Place of Attachment in Human Behavior*. C. M. Parkes and J. Stevenson-Hinde. New York, Basic Books: 185-201.

- McNeill, W. H. (1995). *Keeping Together in Time: Dance and Drill in Human History*. Cambridge, Harvard University Press.
- Merker, B. (2000). Synchronous Chorus and Human Origins. *The Origins of Music*. N. L. Wallin, B. Merker and S. Brown. Cambridge, MA, MIT Press: 315-327.
- National Institutes of Mental Health (1998). Diagnosis and Treatment of Attention Deficit Hyperactivity Disorder (ADHD). NIH Consensus Statement **16**(2): 1-37.
- Quine, W. V. (1960). *Word and Object*. Cambridge, MIT Press.
- Small, C. (1998). *Musicking*. Hanover and London, Wesleyan University Press.
- Sloboda, J. (1985). *The Musical Mind*. Oxford, Oxford University Press.
- Strogatz, S. H. and I. Stewart (1993). "Coupled Oscillators and Biological Synchronization." *Scientific American*(December): 102-109.