The Human Faculty for Music: What's special about it?

John Christopher Bispham[1,2]

Wolfson College

[1] Leverhulme Centre for Human Evolutionary Studies (LCHES)
Division of Biological Anthropology
Department of Archaeology and Anthropology

[2] Centre for Music and Science (CMS)
Faculty of Music

University of Cambridge

PhD Thesis
Submitted 03/2018

Supervisor: Professor Ian Cross

Abstract (short version)

This thesis presents a model of a narrow faculty for music - qualities that are at once universally present and operational in music across cultures whilst also being specific to our species and to the domain of music. The comparative approach taken focuses on core psychological and physiological capabilities that root and enable appropriate engagement with music rather than on their observable physical correlates. Configurations of musical pulse; musical tone; and musical motivation are described as providing a sustained attentional structure for managing personal experience and interpersonal interaction and as offering a continually renewing phenomenological link between the immediate past, the perceptual present and future expectation. Constituent parts of the narrow faculty for music are considered most fundamentally as a potentiating, quasi-architectural framework in which our most central affective and socio-intentional drives are afforded extended time, stability, and a degree of abstraction, intensity, focus and meaning. The author contends, therefore, that music's defining characteristics, specific functionalities and/or situated efficacies are not demarcated in broadly termed "musical" qualities such as melodic contour or rhythm or in those surprisingly elusive "objective facts" of musical structure. Rather they are *solely* the attentional/motivational–frameworks which root our faculty to make and make sense of music. Our generic capacities for culture and the manifold uses of action, gesture, and sound to express and induce emotion; to regulate affective states; to create or reflect meaning; to signify; to ritualize; coordinate; communicate; interrelate; embody; entrain; and/or intentionalize, none of these is assessed as being intrinsically unique to music performance. Music is, instead, viewed as an *ordered* expression of human experience, behaviour, interaction, and vitality, all shaped, shared, given significance, and/or transformed in time. The relevance of this model to topical debates on music and evolution is discussed and the author contends that the perspective offered affords significant implications for our understanding of why music is evidently and remarkably effective in certain settings and in the pursuit of certain social, individual, and therapeutic goals.

Formal Declarations

1) This dissertation is submitted for the degree of Doctor of Philosophy.
2) It is the result of my own work and includes nothing which is the outcome of work done in collaboration.
3) I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.
4) It does not substantially exceed the prescribed word limit of 80'000 words excluding the bibliography, cited references and appendices, for the Degree Committee for Archaeology and Anthropology, and is within limits specified by the examiners in their joint report (+/- 10%).

Related Publications and Presentations

Some of the ideas and arguments in this thesis have been, or are in the process of being, explored and/or advocated in the following media appearances, publications, and conference/seminar talks:

Media Appearances:

- Television Interview with Pascal Goblot (in preparation). Documentary Film on Prehistoric Music. ARTE TV, https://www.arte.tv/en/.

Publications:

- Bispham, J.C. (in press). Music, Evolution and the Experience of Time. Invited Chapter Submission for 'Oxford Handbook of Time in Music' (Oxford University Press).
- Bispham, J.C. (in preparation). Music's Evolutionary "Design Features". To be submitted to Behavioural Brain Sciences (Cambridge University Press).
- Bispham, J.C. (2012). How musical is Man – An evolutionary perspective. Chapter in "Sound Musicianship: Understanding the Crafts of Music. - 1/7 Meaningful Music Making for Life". Cambridge Scholars Publishing.
- Bispham, J. C. (2010). Modelisation de la musique - motivation, pulsation et tonalites musicales. In 'Musique et Evolution' Eds. Deliege, I., Vitouch, O. PSY Mardaga.
- Bispham. J.C. (2009) – Music's "design features": Musical motivation, musical pulse, and musical pitch. Musicae Scientiae, special issue: music and evolution.
- Cross, I., Bispham, J., Himberg, T. &Swaine, J. (unpublished) – Evolution and Musical Rhythm. Submitted to Evolutionary Anthropology.
- Bispham, J.C. (2007) – Music as socio-affective confluential communication? Response to 'a commentary on Bispham' (2006). Music Perception, 25;2
- Bispham. J.C. (2006) - Rhythm in Music: What is it? Who has it? And Why? Journal of Music Perception, special issue on rhythm perception and performance, 24;2, 125-134.
- Bispham, J.C. (2006) - Music means nothing if we don't know what it means - lead review of 'The Singing Neanderthals' by S. Mithen. Journal of Human Evolution, 50, 587-593.
- Bispham, J.C. (2004) – Bridging the Gaps – Music as a Biocultural Phenonmenon. Commentary on 'In time with the music: The concept of entrainment and its significance for ethnomusicology' by Clayton, M., Sager, R., & Will, U. ESEM Counterpoint 1.

Presentations

- Bispham, J. C. (2004) - Music and Evolution. Department of Biological Anthropology Seminar Series, University of Cambridge, UK.
- Bispham, J. C. (2005) - Rhythm in Music: What is it?, Who has it? And Why? Rhythm Perception and Performance Conference II, Ghent, BE.
- Bispham, J. C. (2005) - Evolution, Ethology and Entrainment. Poster Presentation at 'Music and Neuroscience Conference IV', Leipzig, DE.
- Bispham, J. C. (2005) - The Human Skill of Interpersonal Music al Entrainment. International Network of Musical Entrainment Studies Workshop I, Ohio State University, US.
- Bispham, J. C. (2006) - Evolution and Musical Rhythm. International Network of Musical Entrainment Studies Workshop II, Open University, Milton Keynes, UK
- Bispham, J. C. (2007) - Towards a model of musicality: Species-specificity, domain-specificity and universals. Leverhulme Centre for Human Evolutionary Studies Seminar Series, University of Cambridge, UK.
- Bispham, J. C. (2012) - Music's Evolutionary Design Features: Musical Pulse, Musical Tone and Musical Motivation. SEMPRE International Conference on Music , University of Western Australia, Perth, AUS.
- Bispham, J. C. (2016) - Time in Music: Pulse, Pitch and Attentional Architecture. International Conference on 'Time in Music', Faculty of Music, University of Oxford, UK.

Abstract (long version)

INTRODUCTION

The central question and rationales for this thesis are outlined. What is special about music? What is universally present in music across cultures whilst also being distinct from other forms of animal and human behaviour and communication? What constitutes a putative narrow faculty for music? The scope and challenge of this focal question is delineated by four clear statements all of which are in essence, if not in the details, uncontroversial: **1)** Music is, within and across cultural boundaries, enormously and wonderfully diverse; **2)** Most features we commonly regard as being essentially constitutive of "music" and "musicality" – e.g. pulse, rhythm, melody, gesture, vocal affect, performance, expressivity, movement, emotion regulation, synchronisation, harmony – are broadly speaking present in and integral to other forms of interaction and communication; **3)** Music appears to share common substance across cultures in terms of certain consistencies in function, context, setting, and efficacy, as well as in being widely prized and experienced as being intensely meaningful; **4)** The full capacity for music is manifest of a vast interconnected mosaic of skills and abilities ranging upwards from the most basic perceptual, vocal and motoric functions. I argue that, as something like music appears from our cultural vantage point to be recognisable as such across cultures, it is logically pursuant that some identifying features or distinct conglomerations of features are commonly present. A key point here is that these traits may not be physically or knowingly/descriptively evident but may reflect implicitly learned ontology and/or empathy with the underlying psychological mechanisms and experience. Finally I contend that understanding and asserting a theoretical model of a narrow faculty for music is an essential step towards understanding the emergence of music and musical capabilities during the course of hominin evolution as well as for interpreting its modern-day uses, functions and efficacies in social, individual and therapeutic settings. Thus, this thesis aims to contribute to an enhanced applicability of - and clarity in - the ever-growing body of interdisciplinary research into music.

SECTION I – Music and Evolution

Chapter 1 reviews and critiques the current state of research into the evolution of music focusing on debates regarding the status of music as an evolutionary adaptation, putative rationales for the emergence of musical capabilities, universals in music, cross-cultural concerns, and theoretical issues concerning cross-species comparative perspectives. Brief reviews of key literatures on the universality of music in human culture; innate predispositions for musicality; specificities of neurological domains and/or networks for music; paleontological evidence for the emergence of necessary vocal and corporeal capacities; and the archaeology of musical artefacts from the upper Palaeolithic are also presented with particular note to areas of concern to the central arguments of this thesis. This chapter also reviews and explores attempts to position music – biologically and culturally - within broad frameworks of animal and human communication. Following Cross (2007) the notion is described of three concurrently operational dimensions of meaning and affect in music: *motivational-structural* - based on global characteristics of sound with its effects originating in correlates of sound in the physical world and in correlations between signals and the motivational state of an individual; *socio-intentional* - a responsiveness to structural features of human interaction and attributions of intentionality ; and *culturally-enactive* - meaning and affect derived from enculturated, formal and specifically personal learning processes that are rooted in the conventional and institutional use of music. I propose that the the central questions of this thesis be understood to target, in particular, the domain-specificities of the socio-intentional dimension. At an even more fundamental level I posit briefly, following Porges (2003), that the evolution of the mammalian nervous system and anatomy putatively provide a physiological basis for links between visceral state regulation and both the expressive and receptive domains of social communication. Finally definitional issues are considered and, in particular, Cross's (2003) suggested definition that 'music embodies, entrains and transposably intentionalises sound in time and action'. This is seen to be a valuable functional account for an evolutionary framework but also to lack a degree of specificity in terms of describing features that are unique to music.

SECTION II – The Narrow Faculty for Music – Musical Pulse, Musical Tone, and Musical Motivation

Section Introduction - This opening section briefly introduces the work of linguist and anthropologist Charles Hockett on the "design features" of language. This provides a necessary background for the prinicipal methodological inspiration for this thesis, which stems from more recent arguments (Hauser et al., 2002) on identifying a broad and narrow faculty of language – the latter being components specific to language (in a sense unique/distinct "design features"). It then outlines briefly the three main chapters of Section II (chapters 3, 4, and 5) which, on the basis of cross- and intra-species comparative analyses, provide models of musical motivation, musical pulse, and musical tone as species-specific, context-specific, and universal features of an ontogenetically emergent capacity for music. Acknowledging that some musics may lack a salient pitch or rhythmic component, the premise is introduced that music is universally identifiable by the presence of musical motivation and *configurations* of musical pulse and musical tone.

Attention is drawn to two key aspects of the methodological approach taken and their reasoning. The first of these is a concentration on the psychological and attentional mechanisms involved in music rather than on the physical correlates of behaviour. The second point is the attempt to get to the universal core of the question by focusing on those capabilities and processes necessary for engaging appropriately with the ostensibly most "simple" forms of pulse, tone, and motivation. These are tentatively characterised as tapping an isochronous pattern, humming a sustained pitch and an intrinsic desire/impulse to do so. It is noted with reference to both points that music perception is physically embodied and thus no

categorical distinction between perceiving/listening to and producing/performing music is made. Finally each section is informed to then proceed to investigate which (if any) of the mechanisms and capabilities described are unique to music.

Chapter 2 - Musical Pulse - focuses on the psychology and physiology of pulse in music and offers a comprehensive comparative perspective arguing that musical pulse in the most fundamental sense can be identified by the acoustic correlates of engagement with a sustained, volitionally controlled, future-directed attentional pulse. Within a temporal window of the psychological present this incurs awareness of a pulsed framework, is perceived unambiguously or at hierarchically related levels by enculturated individuals, involves engagement of the motor system in a way that enables management of fine & gross temporal control in ballistic and smooth movements, and involves mechanisms for period correction based upon the output of self or others. A key point highlighted in this chapter is the notion that musical pulse is rooted indivisibly in fundamental mechanisms of timing and movement. I propose that period correction mechanisms and a corresponding sustained pulsed attentional framework incurring awareness and degrees of volitional control distinguish music from other 'intrinsic motive pulses' in behavioural and communicative frameworks.

Chapter 3 - Musical Tone - mirrors Chapter 3 in the frequency domain by describing a concept of musical tone built upon the literatures on pitch perception, relative pitch processing, consonance, amusia, the evolution of vocal physiology and neurology in human and non-human primates, volitional control of vocalization, and studies of pitch-shift feedback correction. Also considered are the use of pitch in non-musical interactions and reported incidences of "pitch synchrony" or "tonal synchrony" in linguistic and parent-infant interactions. Musical tone is subsequently described as sustained performance and perception related to stable implicit or explicit psychological pitch centres in which loudness, pitch, and timbre are independently variable. Engagement putatively involves two frequency correction mechanisms - one fast and pre-conscious, the other tied with awareness and partially intentionalisable - based upon the output of self and others and a desire to maintain certain frequency relationships. Similarly to suggestions of chapter 3 I argue that the slower correction mechanisms could be unique to a musical context. This again accords particular importance to music affording a sustained attentional framework, a degree of volitional control and awareness and a constant integration of what has come before with the now and predictions of what is to come.

Chapter 4 - Musical Motivation - presents a discussion on the specificities of an intrinsic motivation for musical action. Debate is initially framed in terms of ethnomusicological discourse on uses and functions of music across cultures. The psychological literature on music and emotion is reviewed and, in particular, the prevailing focus on expression and induction is criticised in not affording a specific, functional and/or evolutionary account of musical emotions. Following Schiavio and colleagues an alternative enactive/dynamic systems approach is preferred. Extending the line of reasoning from chapter 1 and a broad understanding of music as communication, I argue that structural correlates of affect are shared with other human, and in part non-human, communicative behaviours. A suggestion, following Swaine, is therefore presented that music is perhaps better understood as communicating a strategy for interaction and /or the regulation of mood. Critically I suggest that music's particular connections and efficacies with emotion can be directly linked to the specificities of musical pulse and tone described in chapters 2 and 3. I argue, further, that a similar case can be presented for music's effects on other cognitive capacities, social skills, and intersubjective dynamics. Our understanding of the motivations for music, I contend, requires balancing the individual subjective experience with the many interaffective, interattentional, and interintentional drives inherent in interaction and the feelings of shared knowledge, prosociality and synchrony. All are, in music, given a particular space, sustained attentional focus, and extended time within the architecture of musical pulse and tone. Importantly I propose that it the same mechanisms that extend and intensify individual experience that also root our capacity and motivation to synchronise and harmonise in group interaction. This is discussed with particular reference to the universal use of music in ritualistic ceremony. I conclude from the previous three chapters that components of a narrow faculty for music – the specificities of musical motivation, and configurations of musical pulse and musical tone - provide a sustained attentional structure for managing personal experience and interpersonal interaction. They afford a continual phenomenological linking of the immediate past, the perceptual present and future expectation and afford enhanced significance, mnemonic presence, affect, and intersubjective connectedness.

SECTION III – Summary and Conclusion

The concluding section concisely reviews the argument presented, and reminds the reader of its benefits over previous approaches in terms of providing interdisciplinary clarity, and in terms of the potentially fruitful empirically testable hypotheses that are suggested by the model. I argue that the identification of the constitutive components of a narrow faculty for music should act as a focus or constraint on hypotheses regarding the evolution of music and the role of musical capabilities during the course of evolution in the hominin lineage. Two evolutionary implications, in particular, are submitted. Firstly I briefly suggest that those features that appear to be specific to music across cultures could have emerged exaptively from a socially motivated move towards relational processing and from developments in working memory, attentional capacities, consciousness, and the volitional control of action. Music – narrowly defined – may have emerged functionally but non-adaptively as a technology of mind. Secondly I make the point that the capacities for group synchrony and harmonization in music could have first appeared in response to individual selective pressures. Thus we need not necessarily interpret music in reference to group selection models of human evolution and and can avoid many of the associated and embedded problems and controversies. The potential relevance of the findings in section II to topical debates on music's varied roles in modern-day society – its effects on individual well-being; fostering cooperation, affiliation, and group identity, in communicative and neurological therapy; and in education and in community settings - is also touched upon. In particular, I suggest that the specificities of our faculty for music identified in this thesis highlight that future progress in understanding music could come from increased research focus on issues of absorption and attention.

Contents/ Structure

INTRODUCTION

In a world such as ours,... it is necessary to understand why a madrigal by Gesualdo or a Bach Passion, a sitar melody from India or a song from Africa, Berg's Wozzeck or Britten's War Requiem, a Balinese gamelan or a Cantonese opera, or a symphony by Mozart, Beethoven or Mahler may be profoundly necessary for human survival, quite apart from any merit they may have as examples of creativity and technical progress. It is also necessary to explain why, under certain circumstances, a "simple" "folk" song may have more human value than a "complex" symphony.
John Blacking (1973: 116)

While most Western definitions of music stress attributes such as combinations of tones, beauty, intelligibility, and expressiveness and suggest that these attributes are criterial in judging whether something constitutes "music",... there are societies and musics where these criteria make no sense at all.
Bruno Nettl (2005: 18)

Time and measure are to instrumental Music what order and method are to discourse; they break it into proper parts and divisions, by which we are enabled both to remember better what has gone before, and frequently to forsee somewhat of what is to come after: we frequently forsee the return of a period which we know must correspond to another which we remember to have gone before; and according to the saying of an ancient philosopher and musician, the enjoyment of Music arises partly from memory and partly from foresight.
Adam Smith ([1777] 1982: 204)

In a brief scene from the recent historical drama 'The Crown' we see a depiction of the coronation of Queen Elizabeth II in 1953 from the perspective of the former King Edward[1] watching it unfold on the television from a crowded room in Paris. During the procession to the altar, the assembled choirs of Westminster Abbey, St Paul's Cathedral, the Chapel Royal, Saint George's Chapel, Windsor and the Royal School of Church Music, together with a full orchestra, conducted by Sir Adrian Boult, perform Handel's Anthem 'Zadok the Priest'. He offers the following commentary with a brief interjection from a guest:

> Edward: "The oils and oaths, orbs and sceptres, symbol upon symbol. An unfathomable web of arcane mystery and liturgy blurring so many lines no clergyman or historian or lawyer could ever untangle any of it"
>
> Guest: "It's crazy"
>
> Edward: "On the contrary, it's perfectly sane. Who wants transparency when you can have magic, who wants prose when you can have poetry? Pull away the veil and what are you left with? An ordinary young woman with modest ability and little imagination. Wrap her up like that, anoint her with holy oil and hey presto, what do you have? A Goddess!"

In the following scene we cut to Prince Edward, alone in his garden playing a Scottish folk tune on the bagpipes. Here no verbal commentary is offered.

So in a just a short space of time, on one particular day in history, within one complex cultural setting and from just one individual's viewpoint we have already two vastly contrasting manifestations of music. In the former, we witness grandeur, ceremony and global proliferation through the medium of live television on a previously unprecedented scale. The choice of Anthem, composed in 1727 for the coronation of George II and performed at every monarch's coronation since, suggests historical stability and institutional permanence in what is in fact a time of transformation. The few words direct us - the people - decidedly towards rejoicing the occasion while the somewhat incongruent, some might argue paradoxical, conflation of human institution with sublime blessing is given truth in the forceful declaration 'Zadok the priest and Nathan the prophet anointed Solomon King' and again in the following 'May the King live forever, Amen'. The compositional structure is also perfectly suited to the occasion with a calm harmonically driven string introduction of simple chordal arpeggios being followed by a sudden rousing *forte tutti* entrance signaling a period of transition and the arrival of a new and glorious reality. Most of all, however, the coordinated performance of 480 singers and musicians imbues ineffably a sense of scale, unity, and importance. Without needing to explicate the discursively problematic, it engenders appropriate and convergent emotional states in the performers and audience and a sense of collective togetherness – a phenomenological sense of having shared in a meaningful experience. As such the music not only reflects the affective qualia of the event, it is an active, indeed essential, component in the creation of a new national truth.

In the later solitary scene the Prince plays the Scottish military lament 'Flowers of the Forest', an ancient folk tune originally written to commemorate the defeat of the Scottish army of James IV at the Battle of Flodden in September 1513. It was played at the funeral of H.M. Queen Victoria in 1901 after which it gained substantially in popularity. Thus it could, one imagines, have had personal significances as well as associations with issues of regiment, royalty, public occasion and loss. On the day of the coronation these would no doubt have weighed on his mind. However, it seems likely that any symbolic references would have been subconscious or perhaps a secondary motivation at this time. His face is visibly full of genuine,

---

[1] Following his abdication as King in 1936 his formal title was HRH the Duke of Windsor.

powerful and complex emotion that defies simple linguistic categorization. The music seems to afford a space in time for reflection, for processing the events of the day and the consequences of his own complex history and values. It is a form of personal regulation, a therapeutic strategy, an experience that is necessarily beyond the scope of words. Indeed, whereas in the first example language could give coarse pointers and articulate some of what was going on, in this context, it could only interfere, impede attention and diminish the ineffable quality of the experience felt and portrayed.

In both cases it is crucial to appreciate that the musical performance and practice is an active integral part of human experience and reality. It is individually and collectively embodied as well as being socially reflective and generative. Its meaning and significances necessarily emerge and project in symbiosis with culture, the intricacies of interaction and the individual experience from which they are born. As we broaden our view from this individual take on a single day to considering a world full of musics spanning recorded history we find that this observation generalizes profoundly and thoroughly. It applies whether it be to infant-directed singing in the !Kung people of the Kalahari Desert (Konner, 1974) or the Hazda of North Central Tanzania [2](Knight et al, 2003), a Senoi Temiar Shaman's accompanied healing song in peninsular Malaysia (Roseman, 1991), coordinated panpipe playing at a festival in Conima in the Peruvian Andes (Turino 1993), A Karelian lament in Eastern Finland (Tolbert, 2008), or Children's play songs or 'Tshikona' - national dance - of the Venda People in South Africa (Blacking, 1969; 73). In all of these, and the countless and multifarious other uses of music across time and geographical divides, we find music to be functioning in regulating emotional, cognitive and physiological states; mediating between 'self' and 'other'; representing cultural symbolisms; and/or in coordinating individual, dyadic or group actions (see Clayton, 2016).

Music is, within and across cultures, not only active and abundant but also wonderfully diverse ranging from Gesualdo Motets to Bob Dylan, an Indian Classical Alap, Mongolian throat singing and Japanese Taiko drumming to mention but a very few. Indeed, such contrasting physical manifestations of music can appear to reveal little if any definite and observable common ground. Further challenging any generalizable account of music is the fact that neither practical nor analytic musical skills, beyond an appreciation of basic levels of surface structure, translate across divides of cultural style, immersion and learning. It can only ever be appropriately and fully heard, felt and understood with a corresponding practice, knowledge and experience of its *complex histories, values, conventions, institutions and technologies* (See Cross, 2001). If one may perhaps postulate cases where music is seen to be something valued yet separate, commercialisable, and/or purely aesthetic in quality then, at the very least, we can counter that these are extremely rare and afforded by very recent technological advances and cultural quirks. Certainly, they could not provide a suitable basis for a generalizable philosophy and understanding of music as a constituent component of human biology, behavior and culture.

If we widen our lens yet further to more broadly incorporate colloquially termed expressions of "musicality" we find an even more encompassing ubiquity. To avoid potentially contradictory statements we should perhaps understand these more precisely to be widespread features of behavior and interaction that are given particular salience and attentional focus in music. Regardless, it is clear that many, if not most, of music's seemingly integral and most significant features are - perhaps extended or abstracted - but certainly not unique to music. They are seeded in common pre-dispositions and ontogenetic trajectories, manifest in our earliest developmental explorations and are dynamically shaped throughout our lifespans. Fundamentally in movement we are continually creating rhythmic and quasi-melodic shapes, sensing, engendering and expressing our individual and collective vitalities (See Stern 2004). In terms, at least, of semantically describable categories of emotion music seems to share with vocal communication in general a common 'code' or set of cues (Juslin & Laukka, 2003). Physical posture and gesture too are inextricable parts in any fully manifest and successful communication (Kendon, 2004), with joint actions and moments of synchronicity providing empathic attunements and generating points of agreement and accord (Gill, 2016). Mutually negotiated intrinsic motive pulses (Trevarthen, 1999) and tonal synchronies (VanPuyvelde, 2015) provide coordinative frameworks for interpersonal interaction, intersubjectivity, attachment, participatory sense-making (De Jaegher and Di Paolo 2007) and the regulation of affect in parent-infant dyads and beyond (Stern, 1985; 2009). In oratory and/or poetry dynamically generated expectancies and accents can add credence and persuasiveness (Woodall, 1981), while meter and cadence can allow emphasis, subversions of expectations, highlight changes in mood of interpretation and/or aid in mnemonic retention. Additionally, broad analogies of combinatoriality, embedded structure, implicitly acquired "grammars", and recursion in music and language show further overlap both in terms of macro design, organizational principles and the cognitive mechanisms of learning and working memory in particular that afford them (see Rebuschat et al. 2012). This latter point, of course, highlights the need for manifestations of musicality to also be understood in terms of the cognitive and corporeal mechanisms and capabilities from which they arise.

Musicality begins, in accordance with an enactive approach to affect and cognition, at the fundamental levels of embodied sense-making, primordial affectivity, and selfhood; at the origins of our existence as complex bio-cultural beings (see Schiavio et al, 2016). As such it emerges from and motivates the dynamic, relational, and transformative interactions between us as organisms and our physical and social environment. It can be understood to be working as part of a larger ongoing behavioural process of maintaining adaptive, self-sustaining, dynamical stability through communication. Understanding musicality in this way corresponds seamlessly with broad functional models of animal communication in which signals arise organically from organisms' efforts to manage and assess their surroundings. Developing this framework to human communication and music in particular, Ian Cross (2011) posits that music is operational at distinct, yet possibly overlapping dimensions of meaning and affect. Bearing some resemblance to Runciman's (1998) sociologically based proposal for three principal categories of human behaviour – 'evoked', 'acquired' and 'imposed' - Motivational-structural,

---

[2] Knight et al. (2003) argue that these are the two most distantly related human groups currently known.

socio-intentional and culturally-enactive dimensions broadly reflect iconic physical cues, specificities of the human capacity for culture and learned associations respectively (see chapter 1). Significantly this perspective can accommodate the fact that even non-human species are responsive to a selection of generic structural features of music (McDermott & Hauser, 2005) and that, although musics' meanings essentially result from cultural processes and social contexts (*e.g.*, Bohlman, 2000), we nevertheless appear able to engage appropriately with and/or respond to *some* aspects of musics outside of our own culture (*e.g.*, Balkwill & Thompson, 1999; Krumhansl *et al.*, 2000; Nan *et al.*, 2006). It further provides a necessary platform for comparative perspectives on music (see section II), affording an understanding of continuities between animal and human communication and between ancestral forms of communication, modern-day language and music.

Extending our outlook further still to accommodate the fullness of time we see that a full capacity for music is evidently the result of a complex evolutionary history of continuous and cumulative change. Each initially random mutation will have been assembled gradually upon preexisting biology and genetic coding and will have endured for specific reasons and in response to particular environmental and social pressures. Arguments as to whether something akin to music was ever directly – adaptively – selected for will be discussed in chapter 1 and briefly revisited in the conclusion. Regardless, it is demonstrably a critical constituent part of our species and essential to a full understanding of hominin evolution. Despite complex debates on the cognitive and neurological specifics, it is clear that a complete faculty for music is a phylogenetically emergent mosaic of interwoven and mostly shared, or perhaps exclusively shared, subskills (See Foley 2012; 2016). Consequently music is embedded in afferent and efferent connections with more primary processes. Its many regulatory, affective and psychodynamic effects and functions are necessarily contingent upon the whole and the gregiological process of our species' past. Parsimonious cognitive or computer models of music perception and performance risk describing an ostensibly meaningless process if these rooted ancestral connections are not considered. Similarly observations of seemingly musical behaviours in animals can only be understood as being truly analogous to music if similar processes can be shown to underlie the physical correlates of the behavior.

The many issues touched on above will, of course, be expanded on in the main body of this thesis but noticeably, even from this very brief journey opening with two individual instances of music to the broadest possible perspective, we can appreciate that music is both integral to and inextricably interwoven with the full scope of human experience, relationship, history, culture, and biology. This both invites and frames in scope the central questions posed by this thesis: What defines music? Which musical capabilities are domain and/or species-specific?; What is unique about music?; And what is universally present in music? Essentially what constitutes a narrow faculty for music? – those psychological, physiological, behavioural, and/or cultural features that are both universally present across cultures and yet specific to music[3].

Addressing this question has important practical as well as phenomenological and philosophical implications. Although the goal here is not to provide a full definition of music, it is to an important degree a definitional question in that I am seeking to highlight features that are unique to music. A lack of clarity on this issue has led to a situation whereby it is regularly unclear which aspects of an experimentally observed effect, modern-day use or proposed evolutionary functionality are due to specifically 'musical' traits, features particular to an individual cultural manifestation of music, and which are due to other broader communicative or interactive qualities. A viable model of the constituent parts of a narrow faculty for music is an essential step towards clarity, understanding, testable hypotheses and theoretical rationales for the many observed, yet mechanistically poorly understood functions of music. As such it could potentially enhance the applicability of empirical investigations and further insight into how music relates to other human and animal behaviours and broader domains of cognition, communication, affect, motivation, and culture. Furthermore it could perhaps assist our understanding of how diverse approaches and levels of explanation (e.g. Tinbergen, 1963) are linked in cross-disciplinary studies of music. Specific examples that confirm the necessity for clear and decisive definitions and the potential pitfalls of their absence are given throughout chapter 1. The chapter also concludes with discussion and critique of existing attempts to define music (e.g. Cross, 2003).

Debates on neurological, psychological, and behavioural domain-specificity(ies), cross-species investigations, cross-cultural cognition studies and theoretical perspectives on universals are all central to this investigation, to wider evolutionary perspective on music and to definitional concerns. Each of these complex topics has amassed a considerable body of empirical investigation and literature in recent years all of which provides a foundation for the endeavour in this thesis. This may, therefore, appear a somewhat daunting starting point. Intense debate on all these questions has, perhaps unsurprisingly, been resistant to consistent findings, interpretations and opinion. In particular, efforts to explore universals in music across cultures have met with severe scepticism at times. Fortunately, however, there are also reasons to be optimistic. First of all,

---

[3] The term 'narrow faculty for music' takes its cue from highly influential proposals made by Hauser, Chomsky, and Fitch (2002) [HCF] for comparative investigation into the evolution of language and for a distinction between a broad faculty of language – incorporating all the capacities required for language – and a narrow faculty of language, which describes only capacities that are exclusively operational in linguistic processing and/or interaction. The original HCF paper and related literature, and their contextual relevance, are explored fully in Section II Introduction (p. 62-65). However, it is worth noting already that the need and motivation for identifying an analogous narrow faculty for music are equally as strong. Arguments surrounding the evolution of music have been hotly debated, particularly over the last twenty years or so. Unfortunately, the lack of a consensual model of the specificities of music – the narrow faculty for music - remains a source of considerable confusion and a barrier to understanding and communication between researchers. It has often not been clear to what extent researchers are discussing features that are specific to particular instances of music or musical style, features that are more accurately relevant to broader categories of cognition and communication (e.g. constituent parts of a broad faculty for music), or features and/or capacities that are, across cultures, unique to music. Particular examples and further discussion is given in Chapter 1. In that chapter, I will also, repeatedly, make the point that it makes sense to focus discussion on the evolution of music predominantly, albeit certainly not exclusively, on features and capacities that are specific to music.

any effort to address theoretically any of the debates and concerns of the previous paragraph is necessarily dependent upon consideration of the others. The search is, thus, not an exponentially increased challenge. In fact, once specified, constituent parts of a narrow faculty for music are a more focused category. For example, if we accept the notion of cross-culturally valid correlations in music between global characteristics of sound such as high tempo or descending melodic contour with an induced emotional pull towards high arousal and soothing respectively, this could perhaps qualify as a universal or 'statistical universal'. However, they cannot be termed as being in any way specific to music as such relationships – assumedly originating in correlates of sound in the physical world and in correlations between signals and the motivational state of an individual – are widely shared in human and non-human communicative repertoires.

Further ground for confidence comes from the fact that, despite the complexities, commonalities and diversity outlined in the previous paragraphs something that we might identify as music does appear to be share some common substance across cultures. It is built on common biological constraints and a shared generic capacity for culture and is consistently prized and cherished, experienced as being intensely meaningful, and inextricably linked with the expression of emotion and the regulation of affect. It is consistently functional and dependably purposeful in cross-culturally comparable situations and social settings. In particular, it is reliably involved in the management of social uncertainty and dependably features as an integral component of ritualistic ceremonies and markers of social change. Further commonalities include being subjectively experienced as being an "honest" signal; promoting empathy and cooperative behaviour; enabling synchronous group coordination; and engendering a feeling of being in a 'virtual' and 'another's' time – 'in communitas' .

We seem, from our particular cultural vantage point and within our individual linguistic taxonomy, to be confident in our ability to consistently identify a seeming abundance of music as such across cultures. Most researchers, following Blacking (1976), have been confident in ascribing to it the status of being a universal component of human culture. It seems logically pursuant therefore that, as we are reliably recognizing, perhaps preconsciously, a suite of behavioural repertoires as being "music", they share some family resemblance – some common acoustic cues, organizational structure, interactive configurations of features and/or underlying psychological processes. A key point here is that these traits may not be physically or knowingly/descriptively evident but may reflect implicitly learned ontology and/or empathy with the underlying psychological mechanisms and experience.

Following on from this last point I hope to convince that some of the seeming impenetrability of certain questions regarding universalities and comparative appraisals stems from an unhelpful focus on the physical correlates of musical action and on higher-level structural features. We need to note, first of all, that much of what is most critical and/or meaningful is not fully reflected in the acoustically and physically observable products of musical action. In a culture where music can be commercialized as a purely physical entity it may take a leap of faith for many to accept that the material surface is a relatively small piece of the whole and cannot demonstrate the specificities of music. Fundamentally, pitch, pulse and motivation are perceptual and attentional constructs. The latter point is important to appreciate from the outset as it is critical to the central argument in this thesis that music's defining characteristics, specific functionalities and/or situated efficacies are solely the *attentional*-motivational frameworks which root our faculty to make and make sense of music. The reason for describing predominantly an attentional structure lies partly in the approach of focusing on psychological mechanism rather than observable action. More specifically, however, it stems from the notion that attentional control (volitional and non-volitional) is a key mechanism through which we perceive temporally structured events and orient ourselves in a world unfolding through time. According to Jones's (1976) influential and established model we are constantly and dynamically building expectations as to the timing of future events based upon perceived regularities and/or learned patterns. We direct attentional peaks towards these expected moments and in response to violations of expectation. Large and Jones (1990) describe an oscillatory mechanism (see Chapter 2) that, essentially, generates attentional "pulses". Furthermore, these "pulses" are supposed to underpin all forms of perception in time and all dynamic interactions with the physical and/or social world. Broadly speaking, therefore, pulse is a form of attentional structure that is ever-present in tracking the internal and external "rhythms" of life and living. In all our interactions with others it is an essential social affordance and ultimately a key part of our capacity for culture. An attentional pulse, in this sense, is clearly not specific to music. The central question of chapter 2 therefore is not whether a pulse or "communicative pulse" is unique to music but rather if there are specificities in terms of the root architectural structure of attention in music and the mechanisms involved. Relative/beat-based timing, period correction mechanisms, and the sustained, constantly renewing nature of attention in music will feature heavily in addressing this question. It is also critical to note that my position is that this question cannot be satisfactorily approached by an acoustic analysis. A musical pulse is often relatively regular/steady but is not, ultimately, distinguished by any absolute degree of regularity (see Chapter 2).

Broadly speaking, an appropriate participation and sensitivity with a musical style may involve implicitly learned expectancies, vocal physiological attunements, unsubstantiated action tendencies, and/or 'hearing' pulses or harmonic pitch centres that are collectively implied but not physically present or unambiguous. Effectively a significant proportion of the cognitive, physical, even phenomenological experience is beyond a purely observational comparative investigation. In terms of higher-level structural features, by these stages we are already far too entrenched in particular cultural constructs, learned specifics and a wealth of creative possibility to make absolutely generalizable statements. Musical scales offer an interesting example as they have been argued to show cross-cultural constancy in terms using at least five discrete pitches, differently sized steps between consecutive tones, a dominance of small-integer value intervals (Dowling & Harwood, 1986; Schellenberg & Trehub, 1996); and tonal hierarchies (Castellano *et al.*, 1984; Krumhansl, 1990). However, scales are not an absolute requisite for pitch-based interaction to be "musical" and I would argue that the commonalities listed emerge in distinct cultural settings parsimoniously from the fundamental designs of our perceptual and interactional capacities (Chapter 3). Additionally presenting such features as statistical universals in music can risk conveying a somewhat sanitized account of the particular tensions, feelings, and fragrances of specific cultural practice. In order to encapsulate all instances of musical pitch (including, for example, monotone chanting) whilst addressing the issue in terms of its being a nascent capacity for musicality, we need to explore at the root level - the constructive foundations. Similarly to the discussions on pulse in chapter 2, Chapter 3 will focus upon psychophysiological correction mechanisms involved in producing and/or perceiving sustained pitch. These necessarily involve constantly integrating working memories of the recent past, the events of the current moment, and expectations for the immediate future. As such, the concept of an attentional structure is again central to the argument.

The approach taken in this thesis is, therefore, to focus on describing the core psychological, cognitive and physiological mechanisms required to engage appropriately with the ostensibly most “simple” forms of musical pulse, tone, and motivation - characterised somewhat problematically perhaps (see section II introduction) as tapping an isochronous pattern, humming a sustained pitch and an intrinsic desire/impulse to do so. This is, however, certainly not aimed at providing a reductionist account of music. First of all these are merely the architectural groundwork, the framework within which the complete gamut of human experience can be given space (see below). Secondly, even pared down to its putatively most basic building blocks, we find individual and social complexity and interdependence with the whole; with fundamental timing mechanisms, neurodynamic resonances and motoric and vestibular systems (chapter 2), with breath, vocalization and connections to the mammalian autonomic nervous system (chapter 3); and with fundamental human motivations towards shared intersubjective experience and the management of conscious experience and attentional focus (chapter 4).

I further submit that focusing on attentional mechanisms which can be inferred from acoustic correlates but not perceived directly does not contradict identifications of music that rely, somewhat circularly, on our ability to recognise it, or a certain “musicness”. As a species we are empathetically aware, and attuned to others to such an extent that even the seemingly simplest musical task imbues intersubjective, interattentional and interaffective resonance, and an implicit familiarity of human interaction and motivational states. The methodological approach of focusing on psychological mechanisms addresses a thorough representation of musical involvement and is true to the inherent nature of music and musical experience and to how music is integrated and embedded into the fullness of individual and cultural experience. It, furthermore, spans widely across demonstratively active or passive engagements with music dissolving any false dichotomies between “performer” and “audience”.

Building the following argument primarily on the available theoretical models in music psychology and cognition, the principal premise of this thesis is that components of a narrow faculty for music - features that are both universally present and specific to the domain of music - are *exclusively* the motivational and attentional-frameworks which root our faculty to make and make sense of music. Musical pulse; musical tone; and musical motivation are described here as providing a framework for interpersonal interaction and a structure for the individual or social management of attention, motivation and/or affect. In particular, I will highlight the proposal that music is distinguished from other communicative contexts by the sustained, constantly renewed nature of attention and by associated correction mechanisms of pitch and period that seem to incur awareness of change and of the underlying construction.

Of course it is precisely this particular attentional architecture that also enables persons and groups of people to come together in sustained musical interaction and group synchrony. It is tempting, perhaps, to overstate music’s capacity to provide shared subjectivity, an ability to break through dualistic separation between subjective, inner experience and objective, external world. It can appear at times to enable a complete empathic attunement and understanding of how others feel and embody in sound and movement. However, rather than arguing for something so absolute it is more faithful to describe music enactively as providing a *sense* of having shared an experience (see Schiavio 2015b). It is a space in time in which dynamic regulatory negotiations of affect and mood can be played out. It can efficiently inspire a *convergence* of focal point, emphasis, motivation and emotional states without needing necessarily to align completely individual position, past, accounts, interpretation, and/or experience.

Viewing music’s specificity in this way has a range of implications for how we understand the human capacity for music, its character in our many personal theatres and social histories, and for how we provide a truthful and encompassing advocacy for the role of music in education, therapy, and society. Understanding music as a particular suite in wider models of behaviour and communication counteracts any well-meaning, but perhaps naively expressed, notion of the “power of music” or of music having a distinctly direct connection to emotion. Music need not be skilled, aesthetically fashioned, concordant, nor even necessarily particularly pleasant. Bohlman (2000) argues persuasively that music need not be beautiful and indeed that this ascription to it only began to make sense when technologies in the West made it possible to reproduce music as a commodity, a product in which the object, ‘beauty’, could lodge. He goes as far as to suggest even that “the more extensively a group community, or society participates in musical performance, the more beauty might serve to encumber its practices and presence.” (Bohlman, 2000: 3*0*). It is easy also to romanticize – privileging times when music provides personal inspiration, transcendence and/or peak experience - and ignore the observation that music is commonly unremarkable. It generally results directly from the ontological condition of embeddedness. As Bohlman (2000: 30) notes: “(M)usic is so much part of other social practices that there is no need to separate it from them or to attribute special qualities to it”. We should not, furthermore, assume a priori that music is intrinsically something positive (e.g. Brown and Volgsten, 2005). We can perhaps posit that features specific to music inherently structure and predispose it towards social interconnectedness, prosociality and a degree of emotional stability. Ultimately, however, its efficacy needs be understood in terms of alignments between the intrinsic drives, vitalities and intersubjectivities of a particular instantiation and any desirable motivational, emotional and/or therapeutic outcomes.

It follows of course that a full analysis and understanding of any musical action or performance needs to be individually resultant from a complete and balanced interdisciplinary study of the specificities of history, culture, cognition, behavior, psychology and phenomenology. Ideally, semiotic, observational, and scientific perspectives as well as applied practice and subjective knowledge all need to be given equal space and weight. A crucial point that is not often voiced is that the fullest possible understanding actually requires a degree of academic restraint and judgment in neither prioritizing features that lend themselves to neuroscientific exploration or experimental paradigm, nor endeavouring to articulate the ineffable. It is inevitably consistent that the latter can only restrict the scope of the experience and contribute to a bias towards the verbalisable. The temptation to add further philosophical obfuscations can in these instances only widen the divide between the academic and experience. To quote Wittgenstein in his philosophical investigations: "in the end, when one is doing philosophy, one gets to the point where one would just like to emit an inarticulate sound" (1953/2010 p. 99).

Blacking (1995, p. 224-225) succinctly described "Music" as being "both the observable product of intentional human action and a basic mode of thought by which *any* human action may be constituted… its value inseparable from its value as an expression of human experience" (emphasis added). Given the possibility of encapsulating selectively from a full breadth of human experience it seems futile to expect any stable commonalities in terms of the themes and substance given space in music. Music has variously been shown to fit comfortably as principal subject in broad models of communication, interaction, therapy, practice, performance, aesthetic experience, ritual, culture and/or social action. However the extent to which each of these is given prominence can vary extensively. Indeed none, I would argue, are definitively critical to identifying something as being music. Rather than being collectively defined by the interior particulars - the multifariously varied manifestations of action and behavior given time within a musical framework - it is, instead, the specific sustained attentional structures that renovate it into something else, connected yet distinct, mundane yet somehow heightened and meaningful.

Essentially I am arguing that what is definitional and unique about music is that it provides a particular attentional framework for managing individual experience and for structuring dyadic and group interaction. It affords an extended experience - a continual phenomenological linking of the immediate past, the perceptual present and future expectation Features that appear unique to the context of musical engagement do not in and of themselves express or affect anything. Rather they offer a sustainable socio-intentional coordinative strategy – a temporal space in which our our most central human drives towards emotional experience, expression, and regulation; intersubjectivity; social alignment; communicative vitality; cultural belonging; and communitas are afforded a degree of stability and reinforced focus. It can be understood as an educational or therapeutic playground, a protracted, liminal, and enhancing field that allows a memorable event – a ritualization and intensification of emotion and meaning. Music is, in accordance with Blacking's description of it as a primary modelling system of human thought, an indispensable tool of consciousness, a transformative technology of the mind.

SECTION I – Music and Evolution

Chapter 1 – Music and Evolution: A review and critique

There is so much music in the world that it is reasonable to suppose that music, like language and possibly religion, is a species-typical trait of man. Essential physiological and cognitive processes that generate musical composition and performance may even be genetically inherited, and therefore present in almost every human being. (Blacking, 1973 p.7)

As neither the enjoyment nor the capacity of producing musical notes are faculties of the least use to man in reference to his daily habits of life, they must be ranked amongst the most mysterious with which he is endowed. (Darwin, 1874 p. 365)

Interest in the evolution of music and musicality has blossomed in the past 20 years or so and is reflected in an increasingly substantial and widespread presence and focus of attention in academic debates and publications (e.g. Wallin, Merker, & Brown, 2000; Balter, 2004; Cross & Morley, 2007; Vitouch, 2007; Morley, 2013). A key enabling factor of these debates has been the growth and ever-increasing value - both intrinsic and in terms of its relevance to wider investigations - accredited to what can be broadly characterized as the field of music psychology (see Cross, Hallam & Thaut, 2016; Deutsch, 2013). While ethnomusicological ethnographic work (e.g. Titon et al. 2009/2016), has extensively highlighted the fact that musics are extensively a product of, and functional within, cultural settings, studies of, for example, innate predispositions for musical perception (Trehub, 2003; Trehub & Hannon, 2006) or the neuroscientific correlates of musical engagement (e.g. Peretz, 2003; Patel 2010), also suggest a concept of musicality as a biologically grounded capacity. A diverse and growing body of cross-disciplinary research into varied aspects of music and musicality (e.g. Avanzini et al., 2005; Clayton, Herbert, et al. 2003; Cross & Rohrmeier, 2007; Miell et al. 2005; Peretz, 2006) have greatly enhanced our understanding of music and, significantly, have allowed for an understanding of music as a concurrently biological and cultural phenomenon, potentiating a more comprehensive evolutionary perspective (see Cross, 2003). In addition to the intrinsic benefits to musical scholarship, inter-disciplinary research involving music also offers invaluable insights and further understanding in other fields related to human cognition, psychology, sociality, and culture (e.g. Boyd & Silk, 2014; Foley & Lewin, 2013). Put succinctly there is an increasing convergence upon the notion that music and musical capabilities are embedded in the whole of our experience and history - they are key features of who we – *Homo sapiens* – have evolved to be.

A primary motivation for engaging with debates on music and evolution has, of course, been a desire for a more complete understanding of hominin evolution and the pressures – environmental and social - and inherent selective advantages that may have led to the evolution of musical and/or proto-musical capabilities in our species and beyond (e.g. Fitch, 2006; Mithen, 2005). However, evolutionary theory has long ceased to be of interest merely in and of itself. Such is the consensus within academic circles as to the central tenets of Darwin's original hypothesis (Darwin, 1859) and such its widespread applicability and relevance that philosophical debate in almost all fields related to human behavior and psychology appear incomplete without recourse to, at least, the basic principles of evolutionary theory. Even our most complex and discernibly defining behaviours such as language (e.g. Aiello and Dunbar, 1993; Chomsky et al., 2003; Dunbar, 1998; Pinker, 1994 Tallerman, 2005); culture (e.g. Dunbar et al., 1999); intelligence and cognition (e.g. Barkow et al., 1992; Byrne 1995; Carter 1998; Deacon 1997; Dennett 1998); and even spirituality and religion (Alcorta & Sosis, 2005; Atran, 2002; Bulbulia, 2004a; 2004b) have been the subject of heated debate from an evolutionary viewpoint.

Crucially these debates highlight the fact that an evolutionary perspective on features of human behaviour and psychology not only provides philosophical substance but also a framework for integrating multiple perspectives on a given area of investigation. This makes it particularly valuable in terms of understanding music and the current developments in music research. With specific regard to music, Cross (2003b) explains that the great strength of this viewpoint is that "by focusing upon phylogenetic relationships and the attributes that may allow a gene, a behaviour, an organism or an interpersonal/group dynamic to be functional, evolution uniquely offers an integrated framework for comprehending music both biologically and culturally within the wider context of human behavior" (p. 21). This thesis is partially motivated by this theoretical strength and by the need for clarity of understanding and theoretical consensus in interdisciplinary research into music. The focal questions posed here - What defines music? Which musical capabilities are domain and/or species-specific? What is unique about music? And what is universally present in music? - necessarily require a multi-disciplinary approach and the unifying framework that an evolutionary perspective can provide.

However, an evolutionary perspective is more than a practical and unifying decision. It is an essential stage and context in addressing these queries. Understanding the history, the process by which the mosaic capacity for music emerged, is fundamental to appreciating how musical capabilities are embedded in and emerge from broader or more primary mechanisms and behaviours. The point here is that, although we could perhaps attempt to address the question of what is special about music without recourse to evolutionary history, it would inevitably be incomplete, lacking a central thread of enquiry, and would fail to encapsulate fully the complex entrenched nature and qualities of musical experience. Music perception and performance needs to be understood in terms of the many connections that are rooted in the gradual processes and amalgamations of our evolutionary history. Music's efficacies come fundamentally from how our capacities to perceive, make sense of and produce music are structured and from the corresponding associations to broader and more generalizable capacities. Even if, as seems very likely, modern-day uses for music have been transmuted from original functions and the selection pressures that helped to forge the emergence of individual musical capabilities, it is parsimonious to assume that they have not been dissociated entirely in the modern mind and body. Ultimately music has purpose not because of intrinsic physical characteristics of any particular piece or performance but because of who we are, because of our individual and cultural experience and our genetic history. In address to Darwin's quotation at the top of this chapter, it is we rather than music specifically that remain the central mystery. Music as a whole cannot be fully divorced from the sum of human experience, emotion, our communicative repertoires, and the functions and processes of our past. Thus a constant theme throughout this thesis will be that, despite the apparent search for something separate, distinct and distinguishable, neither the whole nor the individual components that make up the complex mosaic that is our faculty for music can ever be analyzed or described fully in isolation.

We need also, at this early stage, to be open to the possibility that none of the individual subskills of a full capacity for music are exclusive to music or that music may, broadly speaking, constitute sites on continua of communicative affective potential, referentiality, ambiguity and intention (See Cross et al, 2013). A vast majority of the individual sub-skills required for a full capacity for music are, of course, in particular shared with pre-linguistic communications with infants and with mature language. It is unsurprising therefore that a number of researchers have argued for shared precursor to both in the form of some kind of protolanguage or enhanced affective communication (Darwin, 1871; Brown, 2000; Mithen, 2005; Morley 2013). There remains however, in my opinion, some considerable confusion in such scenarios as to what might distinguish music from the ancient communicative precursor. Mithen, for example, argues for a precursor common to both music and language in the form of a mode of communication that is Holistic - lacking compositional structure and combinatoriality; Manipulative - concerned with affecting the behaviour of others; Multi-Modal - involving voice and gesture; Musical - making use of rhythm and melody and involving synchronisation and turn-taking; and Mimetic - imitative and involving intentional representations: in short, ‘‘Hmmmmm’’ (Hm5). Unfortunately in his proposal music is never clearly distinguished from Hm5 and is all too often located in contexts, concepts and ideas that are profoundly ethnocentric. The very description of something being "musical" before the advent of "music" seems contradictory and certainly problematic if not unpacked further. One of Mithen's fundamental assumptions seems to be that everything needed for music is present in Hm5. This conflation of identities persists throughout his argument as we are asked to imagine hominids (who are, according to Mithen's argument, utilising Hm5) synchronising vocalisations in a communal song at FxJj50 or confirming social bonds through communal Hm5 singing at Atapuerca. Perhaps even more extreme is Mithen's response to Blacking's statement (1995) that the world would be a better place if we were all able to communicate unselfconsciously through music. ‘‘I agree’’ says Mithen: ‘‘bring back Hmmmmm!’’ (p.271). Others have, notably, been more careful and aware of these potential pitfalls. However, the problem of music's specific distinguishing features in this context and beyond remains to be addressed satisfactorily.

Essentially the literature is lacking, and in want of, a psychologically, physiologically, behaviourally, and culturally grounded conceptualisation of music that both distinguishes and relates it to other human and animal behaviours and broader domains such as cognition, communication, affect, motivation, and culture. Cross-cultural definitional issues and a search for universals have often been dismissed as unrealistic due to the sheer magnitude of cultural variability in music (Ethnomusicology, 1971 vol. 15(3); The World of Music, 1977 vol. 19(2)) or as an unnecessary inconvenience. The first of these points highlights the central and sizeable challenge presented and will be extensively addressed in this thesis. However, the latter point lacks any valid justification. A surprisingly common view in debates on the evolutionary origins of music is that definitions of music are superfluous as ‘we all know what we're talking about’ (and by extension that we all know what everyone else is talking about) and that we are better off spending time on more productive empirical research (e.g. Hauser & McDermott, 2006; Mithen, 2006). A review of the literature and of the surrounding interdisciplinary research, as well as broad debates with colleagues, however, makes it all too clear that this is simply not the case. Equally serious is the fact that a lack of clarity on this issue has led to a situation whereby it is regularly unclear which aspects of an experimentally observed effect or proposed evolutionary functionality are due to specifically ‘musical’ aspects, features particular to an individual cultural manifestation of music, and which are due to other aspects of the engagement . As we shall see throughout this chapter this has been a constant weakness in arguments regarding the evolutionary status of music and its putative functional rationales.

I argue that, rather than taking time away from empirical research, a good and viable model of the specificities of musicality – the narrow faculty for music - can offer clarity, understanding, and testable hypotheses to evolutionary debates and cross-disciplinary studies thus greatly enhancing the efficiency and applicability of empirical investigations. Before moving to the central section (II) where these questions will be addressed directly, this chapter will first selectively review and critique the principal areas of debate in studies of music and evolution focusing on areas of particular relevance and short-comings that result from lacking a consensual model of features that are specific to music.

Music and Evolution – key questions

The question of whether musical abilities have arisen as a result of natural or sexual selection – as an adaptation - or whether they emerged as an indirect consequence of evolutionary pressures towards other human abilities – as an exaptation – has a long, albeit until recently sporadic, history. Spencer (1858) and Darwin (1871) presented opposing views in the literature. Spencer argued that music originated in the prosodic features of emotional speech whereas Darwin, conversely, argued that music preceded language (providing the basis for "impassioned speech") and evolved from animal mating calls by processes of sexual selection. In ‘The Principles of Psychology’ (1890), James argues for a non-adaptionist origin of music, describing it as ‘a mere incidental peculiarity of the nervous system’ (p.417).

Music as a spandrel

In more recent times research into the evolutionary origins of music was given reactionary impetus by Pinker's (1997) description of music as no more than auditory cheesecake - a pseudo-communicative and non-adaptive activity that merely tickles senses and capacities that evolved as a result of other, more survival-oriented, evolutionary pressures, in particular, language. Pinker contends that 'music appears to be a pure pleasure technology, a cocktail of recreational drugs that we ingest through the ear to stimulate a mass of pleasure circuits at once' (p.528). He builds his brief argument with reference to supposed parallels between music and language, auditory scene analysis (e.g. Bregman, 1990); emotional calls (e.g. Darwin, 1872); habitat selection (i.e. environmental sound contours); motor control (Clynes & Walker, 1982; Epstein, 1994; Jackendorff; 1994) and 'something else'. With regard to language he argues that both share, at least the concept of, a hierarchically structured universal grammar (Bernstein, 1976; Chomsky, 1993; Lerdahl & Jackendorff, 1983); the metrical structure of strong and weak beats, the intonation contour of rising and falling pitch, and the hierarchical grouping of phrases. His provocative position on music is that "compared with language, vision, social reasoning, and physical know-how, music could vanish from our species and the rest of our lifestyle would be virtually unchanged" (1997, p.528) and he formulates his challenge to those who may disagree by asking 'what benefit could there be to diverting time and energy to the making of plinking noises, or to feeling sad when no one has died?" (p.528).

The primary problem, however, is that this is quite simply the wrong question. Making plinking noises and feeling sad is not what 'music' as a generic feature of human behaviour is about. Pinker's view of music as little more than patterned sound experienced through audition for hedonistic pleasure, a specialized and passive activity, is severely ethnocentric and at odds with a broad and psychologically grounded view of musical engagement. As we have already highlighted in the introduction, and as will be discussed in greater detail below, music is, instead, something we actively engage in that has a wide range of important and culturally constructed meanings (e.g. Cross 2003a). Of course it is important to point out that, although this completely undermines Pinker's argument, it does not deny the possibility of a non-adaptive origin of music. Another, perhaps more surprising, voice arguing for a non-adaptive origin is music psychologist Bruno Repp (1991) who argues that "there is no reason to believe that there is a universally shared, innate basis for music perception" (p. 259). He continues to argue that "music has not been around long enough to have shaped perceptual mechanisms over thousands of generations. Clearly, music is a cultural artefact, and knowledge about it must be acquired. Moreover, in contrast to speech, this knowledge is acquired relatively slowly and not equally by all individuals of a given culture" (p.260). However, as pointed out by Morley (2003 p.185), this statement is only correct "if we consistently substitute the word "music" with something akin to "the culturally-specific manifestations of musical activity". Repp fails to consider that music cross-culturally may be built upon some universal, ontogenetically emergent, and biologically/psychologically grounded music-specific capabilities (see chapters 2-4) and that evidence for musical activity and the emergence of proto-musical capabilities easily dates far enough into the past for genetic selection to have occurred (see below; Morley, 2003; 2013).

Another form of quasi-non-functional argument for music we touched on earlier is proposed by Mithen (2005) who claims to be countering Pinker's hypothesis primarily on the basis that music is too different from language to be adequately explained as an evolutionary spin-off (p.26) and states, in accordance with the approach taken in this thesis, that ''our ultimate concern is with the evolved physical and psychological propensities that provide the capacities for language and music in our species." (p.11) Mithen's argument, however, supports aspects of Pinker's view in that music is still described as nothing more than entertainment - a relic of a previously adaptive communicative medium that has been rendered useless by the dominance and power of language. As mentioned in the chapter introduction, he argues, following previous musilanguage and protolanguage models (e.g. Darwin 1871; Brown 2000; Wray 1998; 2000), for a precursor common to both in the form of a mode of communication – Hmmmm (see above). Mithen reasons that only mimesis distinguishes this communicative form from communication systems in non-human primates and that with the advent of language Hm5 was no longer required but that ''there remained a compulsion among modern humans to communicate with music'' (p.272). However, one of Mithen's fundamental assumptions seems to be that everything needed for music is present in Hm5. This conflation of identities and often reiterative use of the term 'music' persists throughout the course of his argument. As explained more extensively in Bispham (2006a) it once again appears that the central problem lies in a failure to account for the possibility of features that are specific to and/or definitional of music and musical capabilities across cultures (see below).

It is noteworthy and perhaps indicative of weaknesses in the existing argument, that non-adaptionist claims regarding music and language have also been proposed the other way around. Some authors (e.g. Livingstone, 1973; Vaneechoutte and Skoyles, 1998) have argued, conversely, that language is constructed in part upon features that were originally selected for 'music'. What seems most likely, however, is that music and language share a common foundation in the form of a shared ancestral communicative medium (Brown, 2000; Morley, 2003) with both diverging and emerging for different reasons and possibly at different times in the course of hominin evolution. By now a wide range of evolutionary musilanguage (M) and/or protolanguage (L) platforms have been proposed in the literature (e.g. Wray 1998; 2000 (L); Brown 2000 (M); Huron 2001 (M) Mithen 2005; Tallerman 2007; 2008 (L); Botha 2008 (M); Fitch 2013 (M); see Smith (L) 2012). Without getting into the specific disagreements between the various proposals listed at this stage we can at least conclude, as stated by Huron (2001), that "if music is an evolutionary adaptation, then it is likely to have a complex genesis. Any musical adaptation is likely to be built on several other adaptations that might be described as premusical or protomusical. Moreover the nebulus rubric music may represent several adaptations, and these adaptations may involve complex co-evolutionary patterns with culture. In biological matters, things are rarely straightforward" (p.44).

Music as Adaptation

The assertion that music may have emerged merely as a by-product of adaptations towards other abilities has been strongly opposed (e.g. Cross 1999; Huron, 2001) and many researchers have accepted that the majority view is that “the evidence is enough to suggest that an adaptive model for music should be the default hypothesis” (Foley, quoted in Balter, 2005, p. 1121). An abundance of putative rationales for the evolutionary fitness of music capabilities have been suggested in the literature. These can be broadly categorised as relating to mate attraction (e.g. Darwin, 1871; Miller, 2000; Merker, 2000); coalition signaling (e.g. Hagen & Bryant, 2003; Hagen & Hammerstein, 2007); altriciality and mother-infant interaction (e.g. Dissanayake, 2000, Falk, 2004); cognitive development (Cross 1999; Papousek 1996); and group cohesiveness (e.g. McNeill, 1995; Roederer, 1984; Brown, 2000a; Brown, 2000b; Freeman, 2000; Cross, 2007). Although these have now been extensively reviewed in other publications (e.g. Bispham, 2006b; Cross 2011; 2016; Morley 2013) they represent a necessary background for the thoughts and views expressed in the further course on argumentation in this thesis on the functionalities of music in evolution and modern-day motivations. Therefore, the following section briefly reviews and critiques the main ideas and supporting arguments. It is worth bearing in mind that although, as we shall see, some have considerably more empirical backing than others, these rationales are not necessarily mutually exclusive. It is possible that each may have been more or less relevant at different times in response to different pressures and as such may represent a functional scene in the eclectic story of the development of musical capabilities. Crucially, if some partial truth is inherent in any one scenario it would seem parsimonious to expect that functionalities and connections are still embedded in the full faculty for music today.

Mate attraction

Darwin's own hypothesis for the adaptive strength of music was that it was used as a means of enhancing an individual's reproductive success, stating that “it appears probable that the progenitors of man, either the males or females or both sexes, before acquiring the power of expressing their mutual love in articulate language, endeavoured to charm each other with musical notes and rhythm” (Darwin, 1871, pp. 880). More recently Miller (2000) extended this hypothesis incorporating Zahavi's “handicap” hypothesis (1975; 1997) that indicators have to carry high costs to be reliable, and Fisher's theory (1930) of runaway sexual selection, in which both courtship traits and mate preferences are heritable[4]. Miller's argument is built upon the facts that musical engagement and practice, although seemingly without survival function, is energetically and cognitively demanding and upon anecdotal evidence that American rock, jazz, and classical musicians produce ten times more albums that female musicians, tend to create most productively during ages of peak mating effort, and are often highly promiscuous. Additionally he follows Darwin's assertion (1871) that the vocal organs, particulary in complex displays, are primarily used and perfected in relation to the propagation of the species. The analogy is particularly drawn to bird species where it is well-established that song is predominantly sung solo by males and functions as a courtship display (e.g. Catchpole & Slater, 1995; Martin-Vivaldi et al. 1998; Hofstad et al. 2002). However, this is perhaps the wrong analogy, as non-human evidence for *interactive* vocalisation (i.e. cases in which there exists an observable relationship between the output of two or more individuals) in a subset of 'dueting' avian species (e.g. Faraburgh, 1982; Mann, Marshall, & Slater, 2003; Hall, 2004) and in Gibbons (e.g. Geissmann & Ordeldinger, 1999; Geissmann, 2000) appears to suggest, instead, correlations with monomorphism (in the avian examples), pair bond strength, monogamous social structure, and territoriality (see also Chapter 2). Other aspects of Miller's argumentation also fail to hold up in the face of criticism. The central problem is one of ethnocentrism. A more complete cross-cultural assessment of the hypothesis paints a very different picture (e.g. see Bolt, 2007). Generically music rarely figures in human courtship and hardly at all outside of large cultures (Brown 2000b e.g. Aka Pygmies, Arom & Khalfa, 1998). It is engaged in equally voraciously by men and women (although gender differences, reflecting social structures, in the contextualisation of music are common across cultures) (Bolt, 2007). Further criticisms that have been directed at a sexual selection hypothesis for music include that it lacks specificity and fails to account for the extraordinary diversity of music – contextually and behaviourally. To quote Foley (Balter, 2005, p. 1121) “[i]f it was sexual selection, it would be a lot more restricted. We would see it more in courtship and less in other activities. Musical ability and activity are too widespread”. This does not necessarily imply that sexual selection played no role in the evolution of music. However, any music-specific evidence is to date virtually non-existent and it certainly appears unlikely that sexual selection was, at any point in time, the primary selective force for music. Alternative sexual selection hypotheses have been posited in conjunction with synchronous chorusing (Merker, 2000; Merker et al. 2009) and coalition signaling (Hagen & Bryant, 2003; Hagen & Hammerstein, 2007). These are, however, mechanistically and structurally less analogous than the authors suggest and will be discussed in Chapter 2.

Altriciality and parent-infant interaction

Contrary to the above given hypotheses, Dissanayake (2000) has argued that human music originated in perceptual, behavioural, cognitive, and emotional competencies and sensitivities that developed from primate precursors in survival enhancing affiliative interactions – ritualized packages of sequential vocal, facial, and kinesthetic behaviours – between mothers and infants under six months of age. An increasing period of altriciality appears to have been a continual feature of hominin evolution (e.g. Leakey, 1994; Morgan, 1995; Joffe, 1997) and hence, Dissanayake assumes, resulted in an intense selection pressure for proximate physiological and cognitive mechanisms to ensure longer and better maternal care. She argues that “the solution to this problem was accomplished by coevolution in infants and mothers of rhythmic, temporally patterned, jointly maintained communicative interactions that produced and sustained positive affect – psychological brain states of interest and joy – by displaying and imitating emotions and motivations of affilitation, and thereby sharing,

---

[4] In the absence of conflicting pressures the two can propel each other to extraordinary exaggerations of a given trait (e.g. the peacock tail).

communicating, and reinforcing them" (p. 390). In essence mother-infant interactions are presented as being 'proto-musical' in the sense that they are composed of many elements that are 'musical' such as prosody; rhythmic regularities and varieties; dynamic variation in intensity, rate, and timbre; expectancies; cross-modalities; and a conjoining of movement and sound. Dissanayake lists the benefits of early mother-infant interaction that go beyond physical protection and care as follows:

- They direct and modulate the infant's state or level of attention and arousal; for example, alert, soothe, praise, please (Fernald, 1992: 420).
- They offer emotional regulation and support, thereby assisting the infant to achieve a coherent homeostatic equilibrium (Hofer, 1990).
- They provide acquaintance with the expressive (or prosodic) features of language by which even adults gain important information about others, such as sex, age, mood, and probable intentions (Fernald, 1992; Locke, 1996)
- They give exposure to the prototypical and meaningful sounds and patterns of spoken language (Fernald, 1992).
- They develop cognitive capacities for recognising agency, object, goal and instrumentality, a narrative-like mode of thought and perception or 'protonarrative envelope' (Stern, 1995: p. 92-32), and predispose the infant generally to intellectual and social competence, including intentionality, reciprocity and expansion. Just one note of caution: just as this is not language, it is also not music. Nevertheless, this is undoubtedly a crucial factor/pressure in the move from animal to human forms of communication.
- They reinforce neural structures predisposed for socioemotional functioning (Schore, 94; Trevarthen & Aitken, 95).
- They introduce cultural norms of appropriate behaviour.
- They assist in establishing physiological and emotional dyadic attunement and reciprocity, enabling the pair to anticipate and adjust to each other's individual natures (Beebe et al., 1992) and the foundation for later Bowlbian attachment (p. 393).

Group cohesiveness

The role of music in superstitious or sexual rites, religion, ideological proselytism and military arousal clearly demonstrates the value of music as a means of establishing behavioural coherency in masses of people. (Roederer, 1984)

The view that music functions to enhance group cohesiveness and bonding has been presented diversely by scholars with backgrounds in history, musicology, cultural anthropology, neuroscience, music psychology and ethnomusicology. McNeill's (1995) historically grounded argument, in particular, has been influential in suggesting that 'keeping together in time' (e.g. marching) functions as a means of muscular bonding[5] thus generating feelings of solidarity.

More recently a wealth of studies have supported the positive and functional connections between music, synchronized gesture, and/or joint pulse-based actions and enhanced empathy, social context, prosociality, cooperative behavior, altruism; social perception, connectivity, compassion, non- aggression, and physiological attunement in human adults and children (Mogan et al., 2017; see chapter 4). However, neither these, nor structural or neurohormonal[6] views of music as a social bonding system (e.g. Benzon, 2001), are in themselves sufficient to explain fully music's ability to promote group-level cohesion. Ethnomusicological and/or sociomusicological research, of course, highlights the fact that music very rarely, if ever, stands alone. Generically, it is always about other things, a component of wider social activities and culturally constructed meanings (e.g. Tolbert, 2001). One of its predominant contextualizations in all cultures – ranging from small communities to large industrialized societies (Lomax, 1968) - is that of ritualized group ceremonies that act as markers of culturally important events. This has led many to the conclusion that music is primarily functional in providing a temporal framework, collective emotionality, a feeling of shared experience, and cohesiveness to group activities, thus supporting the efficacy of ritualistic ceremonies and the creation of alternate realities and belief systems (e.g. Roederer, 1984; Bispham, 2006; see also chapter 4). Freeman (2000) theorised that oxytocin release in the brain during group music-making loosens synaptic connections in which prior knowledge is held and thus "clears the path for the acquisition of new understanding through behavioural actions that are shared with others" (p. 411). This view is scientifically somewhat contentious but the underlying idea chimes well with the fact that music is consistently employed in achieving an altered state of consciousness

---

[5] Muscular bonding is a term used by McNeill to describe the kind of social connection that occurs as a result of shared physical action.

[6] In recent literature Robin Dunbar and members of his research team at the University of Oxford have advocated a view that sees group music making as a neurohomonal extension of primate social bonding through physical grooming. Intense and extensive grooming in our primate cousins is thought to be functionally underpinned by the release of endorphins (and most likely other neurohormones) which in turn promotes emotional closeness between the individuals involved (Keverne et al., 1989; Machin & Dunbar, 2011). In humans, however, communities have expanded significantly (e.g. Dunbar, 1998) and it is worth nothing that modern-day hunter-gatherers even periodically aggregate and maintain ties with other groups to form 'mega-bands' numbering several hundred individuals (e.g. Zhou et al., 2005; Hamilton et al., 2007). We effectively operate considerably beyond corresponding models of group size and brain encephalisation in other species (see Aiello & Dunbar, 1993; Dunbar, 2003, 2014) and beyond available time constraints for facilitating bonding through grooming (Dunbar, 1991; 2010; Lehmann et al., 2007).

Launey and colleagues (2016) therefore propose that "synchrony-based activities" (e.g. dance, prayer, marching, music-making) cause the release of neurohormones and as such were developed by hunter-gatherer ancestors to act as "a primal and direct social bonding agent in large-scale human communities, providing an alternative to social bonding mechanisms such as grooming" (p. 780). The research team have supported this hypothesis with evidence showing that synchrony and exertion during dance *independently* raise pain thresholds (Tarr et al., 2015; see also Tarr et al., 2016) and that in singing changes in connectivity and pain threshold occur as a function of group size (Weinstein et al., 2016a; 2016b) (increases in pain thresholds are generally understood as a proximal indicator of endorphin release). Presumably these forms of bonding relate to feelings of group identity and belonging rather than an intraindividual connection (see Pearce, 2014; Pearce et al., 2015; 2017).

(Rouget, 1985; Thayer, 1994; Huron, 2001), in group ceremony, in times of social and political change (Gregory, 1997), and often coincides with totemic representations of cultural knowledge (e.g. Durkheim, 1912). Cross (2001) offers an alternative but complementary approach arguing that music's ubiquity and efficacy in encounters with the numinous are best accounted for by reference to proto-music's polysemy, its "floating intentionality" (see discussion on ‘transposable intentionality’ below). With reference to Sperber’s notion (1996) of relevant mysteries - situations where beliefs or mental representations arise which are contradictory but are each separately related to (and hence relevant in respect of) other mental representations and beliefs – Cross explains that music, like religious ideas, are distinguished from everyday beliefs by their paradoxicality and their relevance, by their broad applicability and their ambiguity. He concludes that “by virtue of these attributes music may thus be particularly appropriate as a means of amplifying, exemplifying or reinforcing in the course of ongoing experience just these attributes of belief that are interpretable as religious; music's indeterminacy may suit it for use as a means of pursuing and perhaps even parsing the numinous” (p.37). Crucially the creation of collective alternate realities or beliefs is prospectively functional at the group level by endowing the actions of others with a degree of predictability (Shennan, 2002).

Clearly one of the principal proximal functions of music is to promote group cooparation, coordination, and cohesion (e.g. Merriam, 1964). Musical engagement is consistently operational in promoting group identity (e.g. Stokes, 1994; Ellis, 1985; Levi, 1994, Nettl, 2005; Negus, 1994); group cognition – the transmission of group history and the planning for group action; group coordination; and group catharsis. However, we should be clear to note that this is not in itself sufficient to propose an ultimate group selection model for music[7]. As observed by Cross & Morley (2009, p. 76) “[t]he potential role of music in selection at the level of the group needs to be assessed in the light of the extensive debate within recent evolutionary thinking on the nature, and existence, of mechanisms of selection at the group level” (e.g. Williams, 1966; Wade, 1978; Wilson & Sober, 1994; Sober & Wilson; 1998; Dugatkin & Reeve, 1994; Durham, 1991; Bowles & Gintis; 1998; Shennan, 2002). This issue will be briefly discussed further in the conclusion of this thesis in light of the descriptions of music’s narrow faculty for music offered.

Music as an Exaptation

Others have remained unconvinced by an adaptive origin of music but have provided more positive, functional arguments that see music as having emerged as an exaptation from the acquisition of other, more primary, abilities. Livingstone and Thompson (2009), for example, accredit the advent of music to the acquisition of theory of mind – the ability to attribute mental states to self and others, whilst Panksepp (2009) views the emotional power of music as being rooted in the evolution of basic motor and emotional systems of the brain. The prior proposal, the importance of theory of mind, as well as its role and minimal requirement in musical performance is relevant to discussion in chapter 4. The latter (Panksepp, 2009), of course, has many implications for the emotional embeddedness of music and, as such, is a useful background for much of the discussion in this thesis. However, it fails also to account for any specificities of music as opposed to emotionality and emotion communication in general.

Music as play

Another argument that is strengthened by increasing periods of infant altriciality in the hominin lineage is that music aids in cognitive and social development. Cross (1999) relates Mithen’s suggestion (1996) that the course of human cognitive evolution is marked by an increase in the capacity to integrate information and skills across domains of individual and innate cognitive competencies with Karmiloff-Smith’s (1992) account of children’s cognitive development in which children, after gradually mastering implicit behaviour in individual cognitive domains, iteratively redescribe these into increasingly explicit and this domain-general terms. He suggests that music or proto-musical behaviours, by virtue of being a multi-domain activity, being flexible to multiple interpretation and also being somewhat ambiguously intentional, may play a functional role in general development - and by implication, in cognitive evolution. Music can thus be understood as a form of functional yet consequence-free play – “a cognitive capacity arising from an infant’s propensities to search for “relevance” in, and mastery over, itself and its world and from early elements in that search, particularly the interactions with the primary caregiver” (p.14). It is effectively a safe space in which social forms and capacities can be explored and in which the development of individual minds and affording structures for their interactions in society can be facilitated. As such Cross argues “it may have contributed to the emergence of one of our most distinguishing features, our cognitive flexibility” and may even “have been the single most important factor enabling the capacities of representational redescription to evolve” (p.31). Ultimately therefore this perspective suggests that our human capacity for culture (see Tomasello, 2010) is necessarily dependent on the capacities that underlie our faculty for music both ontogenetically and phylogenetically and vice versa (see also chapter 4).

---

[7] The only prominent author to explicitly argue for a group selection model for music is Brown (2000a, 2000b). He lists ten points – Group size, Functionality (the context-specific performance of functional song types, e,g. anthems); Both sexes; All ages; Group exclusivity; Collectivist context; Collectivist content; Pitch Blend (e.g. Brown, 2007); Teamwork and coordination in performance; and Meter - to illustrate that performance and cognitive features of music reflect the group nature of human music making. With reference to an earlier article describing a common ancestor to music and language musilanguage (Brown 2000a), argues that pitch blend and metric rhythms are ‘design features’ of music that ‘reflect music’s origins as group selected trait’. Whilst it lacks comparative validation and psychological grounding in Brown’s writing, the last three points of his list mirror the conclusions of chapter 2 and 3 of this thesis that musical tone and musical pulse are generically operational in providing a music-specific coordinative framework for interaction.

Music as attention

Another exaptive possibility for music that has somewhat surprisingly not been given its own space in the literature on rationales for music in evolution is enhanced working memory (EWM) in the hominin lineage (e.g. Wynn & Coolidge 2005; 2007; 2010). Morley (2003; 2013) is a welcome exception arguing that musical instruments themselves are good candidates for the types of technologies that provide evidence for EWM and that 'the act of performance itself, *of formally structured 'pieces' at least*, requires complex feedback between attention and procedural and declarative memory (2013, p. 303 emphasis added). However he also argues that while some specific forms in formalized contexts would require EWM capacities, the "efficacy of musicality and its access to those cognitive foundations do not themselves rely on a EWM capacity" (Ibid, p.303). His contention that less formalized musical behaviours would be possible without EWM contrasts with some of the arguments I will present in the further course of this thesis. Although perhaps referring more specifically to the phonological loop component in models of working memory, the constant linking of the immediate past, perceptual present, and dynamic expectation in all music is seen to be dependent upon skilled working memory and attentional capabilities.

Music as a technology of the mind

Another interesting functional account that views music as having evolutionary significance but in this case without a direct biological/genetic origin is Patel (2008; 2010) who describes music as a "biologically powerful human invention" – a "transformative technology of the mind". Patel uses the example of the invention of fire-making (e.g. Wrangham, 2009) to demonstrate that universality need not imply biological adaptation and interprets the available neuroscientific literature on music as suggesting that music can have lasting beneficial effects on nonmusical, pre-existing brain functions within an individual lifespan rather than originating via natural selection processes. Patel (2008; 2010) argues that the available evidence fails to challenge a null hypothesis that assumes no natural selection for musical abilities. Conversely he argues that the available evidence for linguistic abilities is sufficient to reject a corresponding null hypothesis. Either or both views may well turn out to be correct. However, a comparison of the evidence he gives to support these contrasting views is not entirely convincing. In support of the case for an adaptive origin of language Patel draws on ten principal sources of evidence: Babbling; anatomy of the human vocal tract; vocal learning; precocious learning of linguistic sound structure; critical periods for language acquisition; commonalities of structure and development in spoken and signed language; robustness of language acquisition; adding complexity to impoverished linguistic input; fixation of a language-relevant gene; and a biological cost of failure to acquire language. My reading of the literature, howeverm given in an earlier publication is that

> "the first eight of these sources of evidence are at least equally valid for music or have corresponding parallel evidence in psychological studies of music[8]. Furthermore the specificity of the FOXP2 to linguistic capabilities is not completely clear[9] and I would argue that it is far too early to suggest that there is no biological cost of failure to acquire musical abilities[10]. Crucially, just as it has proved difficult to define "music", it is currently unclear what constitutes a generic failure to acquire musical abilities. Despite some anecdotal evidence for a form of musical "arrhythmia"[11], neurological studies of individuals identified as being "amusic" have almost exclusively dealt with deficits in fine-grained pitch perception (Peretz, 2006; see chapter 3). Many "amusic" individuals are able to engage appropriately with other important aspects of music and retain strong emotional connections to music. Furthermore, it remains to be seen if the results of studies of "amusia" are applicable across cultures (Cross 2010)". (Bispham, 2012 p. 134/5)

The Nature of the Evidence

I have previously argued (Bispham 2006b) that the main statements and assertions in the literature that need to be assessed in evaluating the evolutionary status of music are as follows:

- Every known human society has what trained musicologists would describe as ''music'' (Blacking, 1995 p.224), hence we can expect that music accompanied *Homo sapiens* out of Africa (Cross, 2003a).
- 6-month-old infants are capable listeners (e.g. Trehub, 2003a) and display proto-musical abilities from birth (Papousek, 1996).
- Peretz and colleagues argue for the existence of specialized brain structures for music, with reference to the existence of Amusics who retain emotive and referential linguistic abilities is suggestive (Peretz, 2003; Peretz and Coltheart, 2003).

---

[8] The one possible exception is finding a suitable parallel for sign language. However, I am unaware of any evidence suggesting that the reapplication of a behavioural trait in a different modality is indicative of adaptive evolution. Furthermore it is worth noting that individuals with acquired deafness have continued to perform (for example, the percussionist Evelyn Glennie has been profoundly deaf from the age of 12 [Glennie, 1993]) and compose (e.g. Beethoven) music at very high levels. It is worth clarifying from the outset that music is a intrinsically a multi-modal activity in which sound and movement are intrinsically linked from early developmental stages onwards (Bahrick & Lickliter, 2004; Philips-Silver & Trainor, 2005; Thaut, 2005).

[9] Alcock et al., 2000 have, for example, reported significant detriment to musical rhythmic capabilities in affected members of the "KE" family.

[10] An additional point worth making is that even if this were shown to be the case, it need not necessarily imply that a loss of musical ability would not have been biologically costly at an earlier stage of human evolution.

[11] Also, fascinating anecdotal accounts of cerebro-vascular accidents resulting in a lost interest in music (which is described by subjects as being "flat" or "emotionless") have been reported (e.g. Griffiths et al, 2004).

- Music has been around for a long time. The earliest archaeological evidence for music comes from Geissenklosterle 36,000 years ago (Morley, 2003; D'Errico et al., 2003). We should expect, however, that the use of nonfossilizing instruments (e.g., bamboo, wood, etc.) and the voice preceded considerably any archaeological evidence (see Morley, 2003; 2013).

Notably, in the last decade or so, the questions of infant capabilities and domain-specificity in particular have become vastly more explored and complex and more extensive lists in support of an adaptive status have appeared. Cross and Morley (2009), for example, offer a valuable discussion on this subject. More recently Honing and Ploeger (2014) conclude, following suggestions for a nomological network with eight basic modes of evidence that evolutionary psychologists can use to support their theories (Schmitt and Pilcher, 2004), that there are strong indications that musicality is an adaptation. They present the sources of backing as 1) existing theoretical rationales (Darwin, 1871; Miller, 2000; Dissanayake, 2008; Cross, 2007; Merker, 2000; Patel, 2010; Honing, 2011a); 2) psychological evidence (Trehub, 2003; Trehub & Hannon, 2006; DeNora, 2001; Saarikallio & Erkkila, 2007; Seeger, 1987); 3) medical evidence (Peretz et al 2002; Peretz 2002); 4) physiological evidence (Patel, 2008; Peretz & Coltheart, 2003; Winkler et al., 2009); 5) genetic evidence (Ukkola, Onkamo, Raijas, Karma, & Järvelä, 2009); 6) phylogenetic evidence (Fitch, 2006; Hauser & McDermott, 2003); 7) hunter-gatherer evidence (Conard, Malina, & Münzel, 2009; Rouget, 2004); and 8) cross-cultural evidence (Thompson & Balkwill, 2009; Huron 2009).

I would agree that these all constitute the principal sources of enquiry against which proposals and rationales concerning music's role in human evolution need to be tested. However, while all the individual subject areas can boast some considerable supporting evidence, all are highly complex and none of the can be considered to *fully* indicate preferentially towards an adaptive origin for music (cf. Bispham 2006b). Accordingly, it seems that meeting the determinants for an adaptive trait in terms of identifying corresponding innate constraints and domain-specificities (see Justus and Hustler, 2005) remains beyond our current interpretative scope. It is perhaps surprising, for example, that whilst infant predispositions for musicality are generally regarded as evidence for an adaptive model of music evolution two of the leading figures in this field are currently of the opinion that a comparative perspective on infant musical capabilities suggests that "the most parsimonious interpretation of the available evidence is that infant skills are a product of general perceptual mechanisms that are neither music- nor species-specific" (Trehub & Hannon, 2006, p. 19). They prefer instead the view that the biological significance of music is built upon special-purpose motivational mechanisms (e.g. Merker, 2006) that perpetuate musical behaviours in all human societies with general-purpose mechanisms accounting for the perceptual foundations of music. A key point in this argument is that issues that address subjective values (e.g. infant preferences for consonant vs. dissonant intervals) reflect motivational concerns that may simply not be of relevance to non-human species.

To keep stressing the principal motivation and goals of this thesis, it is unclear to me how these sources could even possibly swing the argument definitively one way or another without a partially definitional model of the putative specificities of the behaviour in question. Rather than extend fully the discussion at this point, indications from psychological, physiological, neurological, phylogenetic, and interdisciplinary sources will be broached as they become pertinent in the course of comparative proposals in chapters 2-4 on musical pulse, tone and motivation. The topic of the evolutionary and proximal functions of music will be also be briefly readdressed in light of the central suppositions of section II on the constituent parts of a narrow faculty for music. Regardless of debates on the direct or more complex exaptive evolutionary origins of music we can safely argue that its emergence was an important and consequential event in our genetic history. The rest of the chapter will therefore briefly discuss the universal pervasiveness of music, the issue of universals in music, innate predispositions for musicality, some theoretical matters regarding domain-specificity, evidence of its long history from archaeology and palaeoanthropology, its position within broad models of animal and human communication, as well as some definitional concerns.

Music as a universal feature of human culture

It seems, by now, almost customary in discussing the universal presence of music to bring up the quotation from ethnomusicologist John Blacking (originally published in 1984) that "every known human society has what trained musicologists would describe as ''music''" (1984/1995, p.224)[12]. This carefully worded phrase has been widely accepted despite the fact that our terminological distinctions are not shared with other cultures. Some societies have terms for individual musical genres but not a generic term for music (e.g. [Keil, 1979]; the Basongye of Zaire, or the Tiv of Nigeria, Navaro [McAllester 1954]), some distinguish between 'recitations' of religious texts or 'devotional song' and 'music' (the Judaic 'ta'ameh ha-mikrah' or the Islamic 'qira'ah' Baily, 2005), while others do not distinguish in their language between music and dance (e.g. Sanskrit sangıta, Thai wai khruu and nkwa of the Igbo of Nigeria) or between music and religion (Candomble Afro-Brazilian religion, Maring of Papua New Guinea at the Kaiko festival)[13]. Trehub and colleagues (2015) quote Stone in noting that "honest observers are hard pressed to find a single indigenous group in Africa that has a term congruent to the usual Western notion of 'music'" (p.7). He argues that the isolation of musical sound from other arts -

---

[12] We should note, however, that although this is a supposition widely accepted in the field of music psychology it is also one that should be continually challenged. Colin Turnbull (1979) writing about the Ik describes a thoroughly depressing account of a society in profound decay, where music, and most other human attributes, such as empathy, does not seem to figure at all. While we can perhaps explain this instance of absence as a symptom of a complete cultural failure, I am aware of at least one recent field investigation which reports an absence of any suite of behaviours that resembly our notion of music in a small and remote forest community. The authors wish not to give details at this stage but it is certainly a possibility we should remain open to.

[13] For a more in-depth discussion of these points from an ethnomusicological standpoint see Nettl 2005.

singing, playing instruments, games, dance, and/or performance - proves to be a Western abstraction. Klaus Wachsmann contends likewise that what ethnomusicologists and others may identify as "music" are more precisely "phenomena that sufficiently resemble the phenomena which I am in the habit of calling music in my home ground" (p. 384). Cross (2010) notes further that the notion that 'music' constitutes a mode of human interaction that is clearly distinct from language also appears absent in some emic cultural conceptions (e.g. Mbendjele hunter-gatherers - Lewis (2009)). It is therefore critical that Blacking's popular quotation is understood in full and within its wider context and intended argument. The philosophical questions inherent in this statement are integral to the point being made and frame much of the further course of argument in this current thesis. The full paragraph reads as follows:

> Although every known human society has what trained musicologists might recognize as music, there are some that have no word for music or whose concept of music has a significance quite different from that generally associated with the word "music" or whose concept of music has a significance quite different from that generally associated with the word "music". Nevertheless "music" is a convenient gloss word that has a similar analytical function to Max Weber's concept of "ideal types". As such "music" can encompass both the enormous range of "musics" which members of different societies categorize as special symbol systems and kinds of social action, and an innate, species-specific set of cognitive and sensory capacities which humans are predisposed for communication and making sense of their environment. "Music" is both the observable product of intentional human action and a basic mode of thought by which any human action may be constituted. The most characteristic and effective embodiment of this mode of thought is what we call "music" but it may also be manifested in other human activities, and even in the organization of verbal ideas, such as Martin Luther King's famous "I have a dream" speech or se of the poetry of Gertrude. (1995, p. 224-225)

We can see that Blacking anticipates a view of music as a universally present biocultural human phenomenon (Cross, 2003b) but also the concept of musicality as an organizational principle that is inherent - albeit perhaps not fully instantiated - in communicatively intentionalised behaviours beyond the context of music. He goes on to discuss the difficulties of defining a given behaviour as being "music".

> To understand "music" as a human capability, as a species-specific set of cognitive and sensory capacities, we must begin by treating the definition of "music" as problematic… but not only must we be able to incorporate in a general theory of "music" the characteristics of all different musical systems, or "musics"; we must also account for the many different ways in which individuals and social groups make sense of what they or someone else regard as "music". (p.225)

He continues to illustrate this point with the example of a Kenyan university student who was, in his opinion unfairly, criticized for producing an "incorrect", albeit entirely logical (in terms of the structural features he chose to stress) analysis of Beethoven's Hammerklavier Sonata and concludes that

> Just as musicologists must take note of alternative readings of the sacred texts, so they, as "music"-making beings, have a right to disagree with indigenous classifications. For example, Muslims do not classify Qur'anic chant as music, but it has features in common with Gregorian and other types of chant which many musicologists regard as "music", and so it can be included in the evidence against which a general theory of "music" must be tested. (p.226)

Furthermore, and in contrast to some writers who view the apparent lack of musical ability in modern Western cultures as evidence that musical ability is somewhat sparse in distribution (Barrow, 1995), Blacking argues that, not only is music present in all cultures, but that musical ability is ubiquitous amongst almost all members within society. In his seminal book How Musical is Man? (1976) Blacking notes that the almost universal distribution of musical competence in African societies suggests that musical ability is a general characteristic of the human species rather than a rare talent[14]. This is not, of course, supposed to suggest that we possess a generic cross-culturally transferable musical ability, but rather than we each are born with inherent predispositions and an ontogenetically emergent capacity for musicality that finds fruition within individual cultural settings. Blacking's example of an obviously fine musician failing another culture's assessment of 'musicality' due to incongruent values and musical concepts anticipates the substantial challenge presented to music psychologists striving to undertake empirical cross-cultural research (e.g. Eurola et. al., 2006). Furthermore, Blacking not only anticipates a view of music as a universal biocultural human phenomenon but also the concept of musicality as an organizational principle that is partially inherent - albeit perhaps not fully instantiated - in communicatively intentionalised behaviours beyond the context of music. Crucially he contends that, whilst respecting that others may note alternatively valid readings of our musics, we have a right to disagree with indigenous classifications and that behaviours that fit our understanding can be included in 'the evidence against which a general theory of "music" must be tested' (p.226).

Before moving on it is worth mentioning that music is an activity that we, as a species, also spend a lot of time and energy on. Huron (2001) supports this point with reference to two diverse cultures arguing that from an ethological perspective

---

[14] Of course, any suggestion that in modern Western culture musical ability is the province merely of the trained few is misplaced. Many people do not consider themselves to be 'musical' and may rarely actively produce music. However, almost all of us have the ability to engage with music in culturally appropriate ways (e.g. dancing at the disco, singing along with hymns at church, listening/engaging emotionally with recorded musical performances). The total lack of ability to engage with music – "amusia" - is extremely rare and even when exhibited may represent a narrower inability to engage with specific cultural manifestations of music (see below).

activities that consume large expenditures of energy, such as primate grooming (e.g. Dunbar, 1997; Spruijt, Hooff, & Gispen, 1992), are more likely to be optimised and have a serious or critical purpose. Huron chooses the examples of the Mekranoti Indians – a group of hunter-gatherers who live in the Amazon rain forest - and the modern United States of America. In Mekranoti Indian culture women lay banana leaves on the ground every morning and evening where they then sit and sing for between one and two hours. The men sing every morning typically around 4.30, but sometimes as early as 1.30am. The men, in particular, employ a highly energetically demanding technique that involves holding their arms in a cradling position, swinging vigorously and singing raucously and with heavy glottal attacks. This continues for up to two hours and is sometimes repeated more briefly at sunset (Werner, 1984). In modern United States of America, also, a large proportion of available resources appear to be allocated towards music. In this case it is money that acts as the best index and, although this is not an ideal ethological measure, Huron (2001) notes that the music industry is substantially larger in monetary terms than the pharmaceutical industry. In truth Huron could have chosen any cultures for his argument as music is generically a crucial facet of human cultural and individual existence, universally being employed at times of personal and cultural transition, during ritualistic ceremony and cultural encounters with the divine (Nettl, 2000), as a means of transforming experience (McAllester, 1971) and in the management of self-identity, interpersonal relationships and mood (Hargreaves & North, 1999).

Innate predispositions for musicality

Infants come into the world extraordinarily well equipped to process their surrounding sound world with extremely fine-tuned resolution of frequency (e.g. Olsho, Schoon et al., 1982; Werner, 1992), timing (e.g. Trehub, Schneider & Henderson, 1995; Werner, Marean, et al., 1992) and timbre (e.g. Clarkson & Clifton, 1995; Trehub et al., 1990). They appear, also, to be equipped with essential generic abilities/predispositions for more specific engagement with musical stimuli. Studies into infant music perception and musical predispositions have been spearheaded by Sandra Trehub and colleagues at the University of Toronto. The reader is referred to her own review papers on the matter (Trehub, 2000; 2003a; 2003b; Trehub & Hannon, 2006) for a full complement to the current discussion that aims to focus on the findings and related discussions that are of most significance to an evolutionary perspective. Specific studies of rhythm, pulse, pitch, and harmonisation in infancy will be discussed in chapters 2 and 3 respectively as they become relevant. Particular attention will be given to evidence of overlap between infant and adult processing abilities suggesting predispositions for musicality and innate constraints on cultural manifestations of music.

Music and domain-specificity

Investigations and studies supporting or countering arguments for domain-specificity or particular neural networks in music will also be appropriately slotted in to the course of discussion in chapters 2-4. However, it is worth briefly covering some of the broad theoretical concerns in approaching the issue of domain-specificity here in order to avoid repeated cautioning in regard to the interpretative scope of the studies mentioned. Critically a search for domain-specificity in music implies, not the search for a full-blown neurally specific or neuroanatomically separable 'music module', but the search for individual components that may reflect unique features of a broad faculty for music which in itself incorporates broad features of human cognition and behaviour (Coltheart, 1999). It is also worth briefly clarifying some central issues regarding specialization, domain-specificity, brain localization, and innateness. In Fodor's (1983) original conception of mental modules these four concepts were explicitly linked as essential and interdependent components of what he termed a mental module and together have been viewed as evidence for adaptive evolution of a given trait (e.g. Tooby & Cosmides. 1992). This latter point, however, has been heavily criticized, amongst others, by Fodor (2000) himself who argues that it unduly ignores debates in evolutionary biology on the differences between adaptation (original purpose) and exaptation (modern use) (e.g. Williams, 1966; Gould & Lewontin, 1979; Gould & Vrba, 1982; Jamieson, 1986). Justus and Hutsler (2005) add that this distinction is perhaps better understood in terms of phenotypic and genotypic changes and they reserve the term “secondary adaptation” for genotypic changes resulting from changes of phenotypic use. The crucial point to make here, however, is that although specialization, domain-specificity, brain localization, and innateness are linked in Fodor's original publication - Modularity of the Mind (1983) - and have often been confounded in debates and the literature, they are dissociable and need to be studied independently (Peretz, 2006). A case in point is Hannon and Trehub's (2006) argument, discussed above, that innateness need not imply domain-specificity and localization. Similarly a special-purpose/domain-specific mechanism for music need not be confined to a focal neural network but may instead employ a wider distributed network with cross-domain and general components and some domain-specific, neurally separable components (Peretz & Coltheart, 2003). Peretz (2006) argues, with reference to studies of face recognition (Farah, Wilson, Drain, & Tanaka, 1998) and voice processing (Belin, Fecteau, & Bedard, 2004) that “localization of such essential modules is possible” but that a review of the current literature indicates that “[p]resently we cannot make similar claims about music” (p. 20). Of course a substantial part of the difficulty is that musical ability and experience are reflected in considerable brain plasticity (see, Muente et al., 2002). For a recent overview of the literature and debates on the neuroscience of music the reader is referred to Peretz & Zatorre, 2003 and Peretz & Zatorre (2005). However, in terms of the current context, the key question is that of domain-specificity. Following Justus and Hutsler (2005) it is evidence for a combination of both innate constraints and domain-specificity that would most clearly suggest that aspects of music were shaped by natural selection. In essence, the issue is complicated by the fact that domain-specificity may occur as a result of general learning mechanisms that can be guided by innate constraints or by an extraction of general principles such as statistical regularities in the environment (e.g. Oram & Cuddy, 1995; Saffran, 2001). Furthermore, Elman et al (1996) argue that domain-specificity may be related to five distinct levels – tasks, behaviours, representations, processing mechanisms, or genes. Hence, even where studies add weight to an argument for an adaptive origin for music, some cautionary restraint is always required in interpreting the evolutionary significance of reported evidence of domain-specificity.

Musical universals

Inherent in Blacking's quotation at the top of this section is the existence of some universally distinguishing characteristics, or a family of resemblance (Wittgenstein 1953), that distinguish "music" from language and other communicative repertoires. The history of investigation into musical universals is marked with brief pockets of interest. A widespread reluctance has, however, been the norm with the dominant opinion, particularly from ethnomusicologists, being that music is too varied and is inter-culturally constructed, interwoven and embedded to such a degree that the pursuit is simply not realistic or constructive. However, in conjunction with the reigniting of interest in the origins of music, Bruno Nettl (2000; 2005) has been an influential voice renewing explicit discussion on universals in the literature. He describes a 'concentric circles' approach with absolute universals in the centre – those features which are present in music at all times and in every moment of musical existence. He is, however cautious saying that, at best, we would be able to describe them as having a "certain "musicness" which we cannot define further except to say that it is distinct from "speechness" and maybe from other sound phenomena". There is, in Nettl's opinion, little in a practical way that we can do with this kind of theoretical abstraction. The next set he posits putatively contains features which occur in all music or in the music of every society. He argues, however that this set is necessarily restricted and invalidated by our incomplete (albeit increasingly substantial) cross-cultural and historical sampling set. He argues therefore that we can at best describe 'statistical universals'. He reasons further that "accepting the idea of statistical universals means abandoning the principle that there is a significant difference between universality and popularity". Brown and Jordania extend Nettl's approach describing 'conserved universals' – features that apply equally to other acoustic processes, like speech or animal song; 'statistical universals' – predominant patterns, 'common patterns' – present in many systems and styles; and 'range universals' - gradient descriptions of categorizing surface features. As the authors acknowledge the term universals is a misnomer. More precisely their approach represents cross-cultural trends or typological generalizations (Good, 2008). Steeped in an interest in (re)-invigorating the discipline of comparative musicology what they describe is akin to an updated proposal for cantometric classification and will perhaps be of more considerable value as such.

Nettl is of course correct to point out our ignorance of the many musics present and many more past. Nevertheless, we can perhaps be more optimistic in approaching absolute universals – features common to all phenomena we hear as "music" - if, instead of comparing observable correlates of engagement, we focus on underlying mechanisms. Other authors have variously structured their arguments, at least in part, upon psychological systems and on innate predispositions. Stevens and Bryon, for example, highlight auditory grouping and segmentation, perceptual and cognitive constraints, hierarchical organization and relational processing, implicit learning/knowledge, temporal entrainment and multimodal integration. My proposed approach - written with the intent also to identify features specific to music (see Section II) - goes further in suggesting that music's defining characteristics are *solely* the attentional frameworks which root our faculty to make and make sense of music. This has some notable advantages over the long list of components itemized in Brown and Jordania's categorical structure. For example, they describe an association with dance as a 'common pattern'. If we, instead, consider music more broadly to be rooted in dynamic attentional mechanisms that are inextricably linked with our motor systems, we not only avoid problematizing on the indissolubility of music and dance, but we are, arguably, closer to describing the vitalities of engaging with music and we more fully encapsulate the breadth of musics and musical experience world-wide.

Music as a species-specific trait

Cross-species comparative perspectives are, of course, a crucial component in any evolutionary debate. Experimental studies (particularly with primate species) and some behavioural accounts will form a crucial part of the argumentation throughout Section II. Hence, as with domain-specificity and infant dispositions above it is worth briefly introducing some broad methodological concerns. Tecumseh Fitch sets out an agenda for and a review of the field of biomusicology advocating four foundational principles: Musicality encompassing multiple components; explanatory pluralism – the need to consider all of Tinbergen's explanatory levels; a comparative principle embracing homology and analogy; and an ecological priniciple seeking broad ecological validity. He then structures his review on 'four core components of musicality: 'Song': complex, learned vocalizations; Instrumental music: percussion and drumming; Social synchronization: entrainment, duets and choruses; Dance: a core component of musicality. Selected studies for each component, and studies on other areas of comparative cognition, will be discussed in Section II, but it worth noting now that, by focusing on wide inclusive definitions of each component and on observable phenomena, Fitch casts a wide net and provides a stark contrast to the focus of this thesis. His approach has some definite theoretical and practical advantages. In a scientific environment predicated on that which can be seen, measured and objectively described, it is a prudent method. However, the distinction between observable behaviour and underlying mechanism is important in comparative music cognition too, with significant implications for how we evaluate cross-species studies. As submitted above (see Introduction), there can be a considerable difference between seemingly similar observed phenomena and genuine psychological and structural overlap resulting from a parallel underlying mechanism. How an act is produced, perceived, and processed can, at times, be as functionally significant as the resultant behavioural correlate. Indeed, this seems to be particularly the case with aspects of music cognition and music-making. At all levels of musical structure its effects, implications, values, and resultant transformations are linked, psychodynamically, to how sounds and actions are perceptually ordered, constructed and perceived. Fitch acknowledges this point, to some extent, by advocating explanatory pluralism. Nevertheless there remains a theoretical discussion to be had whether or not observations of components of musicality in nature can correctly be regarded as behavioural or functional homologies or analogies without a full understanding of the attentional and motivational systems and implications.

Music's long history – archaeology, genetics, neurology, and physiology

The previous section established music's ubiquity across cultures and that it can, from our cultural lookout, be distinctly identified. Music's long history is also evidenced in the archaeological record, and in sources from genetics, neurology and paleoanthropology. A brief summary is a necessary background for an understanding of the methodological rationale in this thesis and the often referenced embedded nature of musical capabilities and experience. However, for a much more extensive exposition and discussion on these issue readers are referred to Iain Morley's thesis (2003) and book (2013) on the subject which provide a much more thorough account.

Archaeological evidence for music

Considerable interest in the archaeology of music resulted from the finding of the proposed "Neanderthal flute" from Divje Babe Cave in Slovenia (Fink, 1996; Turk, 1997; Turk et al., 1997). This bone "flute" was found in the middle Paleolithic layers of the cave and was hence thought represent the world's oldest musical instrument. However is has since been persuasively argued that the holes on the bone are more likely to have resulted from carnivore damage (d'Errico, 1998; d'Errico et al. 1998a; Chase & Nowell, 1998). Upper Paleolithic sites, and particularly sites in Isturitz France (e.g. Lawson & d'Errico, 2002), and Geissenkloesterle, are, however, replete with clear evidence of complex musical instrumentation. Perhaps most unequivocal are bone flutes/pipes but we also find evidence for pierced phalanges - "phalangeal whistles"; bullroarers (Dams, 1985; Scothern, 1992); bones with parallel grooves – "rasps" (Huyge, 1990). Even caves have been interpreted to represent "musical" spaces (Glory, 1964, 1965; Dams, 1984; 1985; Reznikoff & Duavois, 1988; Dauvois, 1989). For a full review of findings from the upper Paleolithic the reader is referred to Morley, 2013; 2002 (chapter 3) and d'Errico et al. (2003). Reasonably assured evidence comes in the form of flutes/pipes from the Hohle Fels Cave and Geissenkloesterle sites in Germany. Findings from the latter have been dated to 36'800 +/- 1000 years BC and were found in context with Aurignacian II split bone points (Hahn and Münzel, 1995; Turk and Kavur, 1997; Richter et al., 2000). Even older still are a largely complete flute and two small flute fragments that were found in the oldest Aurignacian layer at Hohle Fels Cave in the Swabian Jura (Conard et al., 2009). The flute is made from bird bone, and dates from the early Aurignacian, 40,000 years ago. Crucially it must be assumed that musical behaviour and the use of musical instrumentation considerably predates these findings. First of all the flutes are highly complex in design (Scothern, 1992) and are hence unlikely to represent the first instance of material objects for musical engagement. Furthermore, ancient societies would, as is the case in modern-day hunter-gatherer societies, most likely have produced music primarily vocally and percussively employing biodegradable materials (see Morley, 2002). Experimental and methodological work on the potential sound-producing properties and analyzable traces of intentional sound production on lithic archaeological materials has also been argued to indicate instrumental behaviour in prehistory (Cross et al., 2002; Blake, 2006).

Music out of Africa

It appears safe to suggest that musical behaviour is a considerably ancient human capacity. As a universally present behaviour (see above) we can expect that musical capabilities were present circa 150,000-287,000yrs ago in the 6900 or so individuals ancestral to the modern-day world population (e.g. Reich & Goldstein, 1998) and certainly at the time of the last dispersal out of Africa circa 60-80,000 yrs ago (e.g. Lahr & Foley, 1994). The molecular data strongly support a common origin for all extant humans somewhere around 100,000 – 200,000 years ago (Ayala & Escalante, 1996; Wood, 1997; Bergström et al., 1998; Cavalli-Sforza & Feldman, 2003). The so-called 'Mitochondrial Eve' (Cann et al., 1987; Saville et al., 1998; Cavalli-Sforza, 1998), the putative common ancestress of all modern humans, was at the forefront of this molecular wave. She has since been joined by a corresponding 'Y-chromosome Adam' (Fu et al., 1996; Dorit et al., 1995; Pääbo, 1995), as well as by data from non-sex-linked genes (Fischman, 1996), and from X-chromosomes (Disotell, 1999)[15]. Following Johansson (2005) it follows that the origins of biologically grounded behaviours such as language and music can almost certainly not be more recent than 100,000 years ago.

Prehistory of musical capabilities

A full perspective on the prehistory of music necessarily incorporates evidence on the palaeoanthropology of vocal control and perception, neurological evidence, vocal physiology, the evolution of breath control, and auditory perception. Once again specific details of some of these arguments will be evidenced at appropriate points in the main section of this thesis (chapters 2 and 3). It is worth noting, though, Morley's (2002; 2013) conclusion that the paleoanthropological evidence supports the view that adaptations without which musical engagement would not have been possible, start to be exhibited in *Homo ergaster* 1.75 Mya. He concludes that from the time of *Homo heidelbergensis*, hominins in the lineage leading to modern-day humans would, most likely, have possessed modern vocal and perceptual versatility and would have been able to produce complex, varied, and planned vocal sequences. Of course, this need not imply that they were engaging musically. The nature of the paleoanthropological evidence cannot possibly distinguish whether or not these capabilities were instrumental in a musical form, propositionally intentionalised or used in some other form of communicative medium. In fact, most of the sources discussed were first proposed as supportive evidence in discussions of the evolution of language. However, the evidence does strongly confirm that at least these components of the modern capacity for musicality have a very long evolutionary history and that capabilities employed conglomeratively in modern musical engagement evolved, at different times and for different reasons, as a complex mosaic (Foley, 2004). Whatever the transitional stages may have been it seems almost certain, especially in conjunction with evidence for a complementary co-evolution of vocal and

---

[15] See Jobling et al., (2013) for further discussion on human origins and evolutionary genetics.

auditory capacities (see chapter 3), that these capacities were in service of pressures towards increased sociality and were primarily communicative in function.

Music as Communication

An evolutionary perspective on music necessarily is predicated on and requires a view that music emerged operationally as a particular strand in the evolution of communication. Music need not be “art”, nor need it even be particularly beautiful or aesthetically attractive. It is overwhelmingly not autonomous (or “pure”) and cannot be separated from the whole of our individual or collective experience and consciousness. Outside the modern and industrialised Western world in which music can sometimes appear as a commodity – a commercialisable physical object – it can have no independent existence or origination. Music is fundamentally an active and shared social behaviour (e.g. Small, 1998) - a communicative act.

As such music has shared fundamental origins in ancient mammalian evolution, in important phylogenetic changes in the neural structures regulating the autonomic nervous system. Porges, (2001; 2003; 2007) proposes a polyvagal theory in which the evolution of the mammalian autonomic nervous system provides the neurophysiological substrates for the emotional experiences and affective processes that are major components of social behaviour. A crucial part of his model is that the area of the brainstem that regulates the heart (i.e., via the mammalian or myelinated vagus) also regulates the muscles of the head, including those of the face, middle ear, mouth, larynx and pharynx. Collectively, these muscles function as an integrated ‘Social Engagement System’ that controls looking, listening, vocalizing and facial gesturing. Effectively he posits that there exist direct connections between control of the sympathetic nervous system via a vagal brake on the heart and our primary means of communication. Effectively mammalian communication is inextricably embedded in an evolutionarily ancient functional origin in physiological regulation[16]. Evidence of the continued relevance of this model, despite some inconsistent findings and questions of interpretation, to music today can be found in studies showing effects of music on heart rate variability (e.g. Ellis & Thayer, 2010; Manzano et al, 2010; Nakahara et al, 2009; Zhou et al, 2010; Krantz et al, 2011; Riganello et al 2010; Orman 2010; VanPuyvelde et al., 2014).

In general it is also important to note and understand that communication does not principally denote a directional conveyance of a propositional ‘message’ or emotional ‘expression’. Broadly understood communication can be viewed functionally as being part of something beyond oneself, as part of how we engage and enact dynamically with our environment (see Schiavio et al 2014; see chapter 4). Indeed models describe animal communication as being principally directed to managing and assessing one’s physical environment (Owings & Morton, 2004; Seyfarth and Cheney. 2003) and human communication as being additionally operational in the social, intersubjective and even the mentally projected realities of our interactional and cultural - ‘virtual’ - truths.

Cross (2007a; 2007b) extends these models in positing the generic existence of three dimensions in music: A motivational-structural dimension based on global structural features such as rate, loudness, pitch and intensity with affect and meaning originating in the fundamental nature of sound in the physical world and in correlations between signals and the motivational state of an individual; a socio-intentional dimension describing specificities of human interaction that derive from partially innate, partially learned sensitivities to structural features of human interaction incorporating attributions of intentionality – i.e. ‘narrative interpretations’ (Bruner, 1986); and a culturally-enactive dimension which encapsulates meaning and affect derived from enculturative, formal and specifically personal learning processes that are rooted in the conventional and institutional use of music as well as the frequency of co-occurrence of music and personal or social situations. This proposed framework bears some resemblance to Runciman’s (1998) sociologically based proposal for three principal categories of human behaviour – ‘evoked’, ‘acquired’ and ‘imposed’ – as well as to proposals made specifically on the roots of emotional experience in music[17]. The framework suggested by Cross (2007a), however, can be considered different from previous schemes in that the dimensions are not considered to be types of behaviour (as in Runciman’s model) and in that “the dimensions of meaning in music are intended to separate out the properties of music as experienced that bear on both emotion **and** meaning and to ground these within a biologically and culturally viable framework” (*ibid* p.18). As such this framework offers a crucial quasi-reference point for the comparative perspectives taken in this thesis (see below and Section II) and is, hence, worth explicating and exploring in some more detail.

The motivational-structural dimension - The concept (and the term itself) of a motivational-structural dimension of meaning and affect in music follows from Owings and Morten’s (1998) model of animal vocal communication in which communication is thought to occur as a result of individuals’ attempts to ‘manage’ and ‘assess’[18] their environment based upon biologically relevant information. A key component of the model is that global structural features of vocalisations can be seen to be indicative of the physical nature and motivational state of the animal. In other words an ‘assessor’ of a signal that is low pitched, has a fast rate, and increasing loudness may do well to be predisposed to ‘assess’/react to this in a manner that would be appropriate of action when faced with a large, highly aroused animal that is getting ever closer – i.e. most likely run away, fast! Equally a social animal faced with a dangerous situation could be expected to benefit from managing

---

[16] This origin in regulatory function has particular relevance to the discussion in chapter 4 on music as a form of emotion regulation.

[17] Sloboda and Juslin (2001) describe music emotional connotations and affect in terms of features that are ‘iconic’, ‘intrinsic’ (based on specifically ‘musical’ features e.g. expectancy and tension [e.g. Meyer, 1956]) and ‘associative’. Lavy (2001) also prefigures the approach taken by Cross in describing the emotional reaction of classical music listeners in terms of music processed as sound, utterance, context and narrative.

[18] The terms ‘send’ and ‘receive’ (Seyfarth & Cheney, 2003) are similarly offered but connote, most likely unintentionally, a less functional account.

the behaviour of conspecifics and/or kin by emitting high frequency, high rate signals that act to direct their attention and heighten arousal levels. Indeed this is exactly what has been shown to occur in Californian ground squirrels (Owings & Hennessy, 1984); marmots (Blumstein & Armitage, 1997); vervet monkey alarm calls (e.g. Owren & Bernacki, 1988); macaque "barks" (e.g. Fischer, 1998); and baboon "lost" calls (e.g. Rendall et al., 2000). Crucially this view of communication need not imply any form of intentionality, reciprocity, or propositionality, but simply implies that animals are predisposed to react in certain ways in certain situations and/or to stimuli and that their predispositions in turn become part of the physical environment of others. As such the management and assessment mechanisms of a group of 'interacting' species may be subject to dynamic evolutionary pressures to exploit each other's communicative and behavioural repertoires. In other words conflicts between the interests of managers and assessors are to be expected and signals need not be directly or verifiably expressive of physical structure or internal states. Instead it is more precise to suggest that communication evolves dynamically *from* direct structural-motivational correlations and that these act as a constraint on communicative repertoires and levels of communicative 'dishonesty' (see Fitch & Hauser, 1995). A common and widespread example of both 'dishonesty' and biological constraints are exaggerated expressive size symbolism and signalling of willingness to fight in potential conflict situations (e.g. Ryan & Brenowitz, 1985; Enquist, 1985; see Bradbury & Vehrencamp, 1998; Johnstone, 1997). A good practical example of the dynamic nature of Owing's and Morten's model – as well the potential complexities that can emerge from its basic principles - is that of male frogs' (*Psysalaemus pustulosus*) mating signals in Panama. These frogs face contrasting selective pressures to exploit female mating preferences for low 'chuck' sounds (Capranica, 1977; Ryan et al., 2001) and to avoid bats that, in turn, have putatively adaptive sensitivities to frog 'chuck' sounds. Male frogs thus balance 'chuck' and 'whine' sounds depending on the proximity of females and/or bats producing what appear to be rather complex vocal displays (Ryan, 1980; 1985; Rand & Ryan, 1981).Cross (2007a, p.6) notes that "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. Hence in terms of music as sound a close relationship can be postulated between the motivational states of listeners and the global structural characteristics of musical sound"[19].

Empirical studies have - despite considerable variation in the experimental paradigms employed[20] - consistently supported this supposition showing that global characteristics of sound in music such as tempo, loudness and pitch are correlated with emotion recognition and elicitation both in terms of categorical (e.g. Ekman, 1992; Oatley, 1992; Plutchik, 1994; Power & Dagleish, 1997) and dimensional (e.g. Russell, 1980; Thayer, 1986) approaches to emotion (see Gabrielsson & Lindstrom, 2001; Juslin, 2001 for reviews; Husain et al., 2002; Cross & Kronhaus, in preparation). Significantly, the most unequivocal correlations appear to be related to arousal/energy levels and, to a slightly lesser degree, valence. Additionally psychophysiological cues such as tempo, complexity levels, and loudness appear to generalise across cultures (Balkwill & Thompson, 1999)[21]. Evidence reported by Huron and colleagues (2006), for example, adds weight to the suggestion that music incorporates biologically significant signals. Following cross-cultural and ethologically grounded suggestions by Bolinger (1964), Ohala (1984) and Morton, (1994) they report that higher-pitched non-vocal melodies are generally interpreted to be less 'threatening' than lower-pitched equivalents.

The above-mentioned studies in themselves present a considerable weight of evidence in support of an operational motivational-structural dimension in music. Nevertheless it is important to note the motivational-structural model primarily describes a *functional* code. Hence it is not surprising that the clearest evidence for its operation in music comes from correlations of form and functional context that are generalisable across cultures, *and* across the full range of human and non-human communicative behaviours. It is well established that 'melodic' contours (direction of F0) are consistently and cross-culturally functional in affecting conspecifics' arousal levels in adult human vocal communication (Scherer, 1986; Bachorowski & Owren, 1995; Scherer et al., 2003); 'intuitive parenting' (Papousek et al., 1991; Papousek, 1996); professional animal handlers (McConnell, 1992); and in monkey communication (Newman & Goedeking, 1992). These correlations appear to generalise to music and in particular to cross-culturally manifest 'functional song types' such as infant play songs (e.g. Trainor, 1996; Rock et al., 1999) and/or lullabies. The latter have been particularly well studied and have been shown to be cross-culturally distinguishable – both by adults and infants – from other types of song purely on the basis of structural features of the music such as low levels of complexity and a predominance of descending intervals[22] (Unyk et al., 1992; Trehub et al., 1993a; Trehub, 1993b). Additionally, and in full accordance with a functional interpretation, infant-absent[23] and infant-present recordings are distinguishable cross-culturally by infants and mothers on the basis of physical acoustic measures (Trainor et al., 1997)[24]. Salient 'performative' cues include increased 'jitter', more 'resonant' timbres, slower tempi, and more pronounced structural boundaries in infant-present versions. Crucially in terms of the current discussion structural and performative cues mirror those in other similarly functional parent-child (Fernald et al., 1984; Papousek et al., 1985) and adult interactive contexts (Fonagy, 1978; Trainor et al., 2000). Structural features such as

---

[19] Cross argues in support of this that it is global structural features of film musics (e.g. Jaws, Psycho) that account for a large portion of its affective functionality (Cohen, 2001).

[20] Gabrielsson and Lindstrom (2001) describe five main methodological categories that have been employed: I. Free descriptions and choice among descriptive terms (e.g. Gilman, 1891; Gundlach, 1935; Rigg, 1937; Watson, 1942); II. Rating and Multivariate analysis techniques (e.g. Kleinen, 1968; Wedin, 1969; Gabrielsson, 1973; Nielson & Cesarec, 1981); III. Continuous recording of emotional expression (Krumhansl, 1997; Schubert, 1999; 2001); IV. Non-verbal responses (e.g. Clynes, 1977; Clynes & Walker, 1982; DeVries, 1991); V. Specially composed music (e.g. Thompson & Robitaille, 1992).

[21] Interestingly in this study pitch levels did not contribute to cross-culturally consistent correlations. This may be related to a music-specific 'production code' (following Gussenhoven, 2002) and will be discussed again in chapters 3 and 4.

[22] The only non-significant (although suggestive) result was in comparison to love songs (Trehub et al., 1998).

[23] Significantly singers were instructed to sing 'appropriately' and with 'feeling' so as to minimize any deliberate intentional biases.

[24] Both mothers and infants predictably showed a preference for infant-present versions (Trainor, 1996).

descending frequency patterns and a slow rate are even generalisably associated with low arousal cross-taxa (e.g. McDonnell, 1991). The evidence, in short, strongly suggests that aspects of music are functionally rooted in biologically grounded biases and processes of a motivational-structural communicative code[25].

The socio-intentional dimension - Despite the clear support given to the notion of an operational motivational-structural dimension of affect and meaning in music, it is equally clear, for example, in the situated performance of a lullaby that further operational aspects of the interaction go beyond the confines of animal communication models. Most crucially a mother's performance is largely initialised, built upon, and affected by inferences regarding the mental and physiological state of the infant. One way of viewing this would be to argue that this reflects an extended human-specific motivational-structural code in which inferential awareness and a desire to manage not only the physical and biological environment, but also the cognitive environment of others (e.g. Sperber & Wilson, 1986) becomes part of an individual's motivational state. However, this would perhaps not capture the magnitude of importance that should be attached to this fundamental distinction. Cross therefore argues for a second dimension of meaning and affect – the socio-intentional - in human interaction. This he argues "is rooted in shared, generic, yet solely human, characteristics of our engagement with the world and particularly with each other". He continues with specific regard to music that "this dimension [is] oriented towards attributions and interpretations of intention and intentionality in engagement with music" (p. 8). Importantly, Cross views the socio-intentional dimension as being rooted in a generic human capacity for culture. To quote Sperber and Hirshfeld (1999) "it is generally agreed among cognitive and social scientists that cultural variation is the effect, not of biological variation, but of a common biological, and more specifically cognitive endowment that, given different historical and ecological conditions, makes this variability possible" (p. 115). Despite debates as to definitions of culture[26] and evidence for the presence of certain cultural phenomena in non-human primates (e.g. McGrew, 1998; Whiten, et al., 2001) it seems, following Cross (*ibid*.), that a human capacity for culture is, uniquely, evidenced in "*shared ways of understanding the world, of understanding each other as ourselves, and of acting together in and on the world, that are transmissible by non-genetic means"* (p. 3., italics in original). This quotation derives its content from an influential paper by Tomasello and colleagues (2005) in which they "propose that the crucial difference between human cognition and that of other species is the ability to participate with others in collaborative activities with shared goals and intentions: shared intentionality" (abstract). This appears, on the basis of a considerable body of comparative evidence, to be a more precise formulation than previous suggestions regarding the specific nature of a human cultural capacity and theory of mind (e.g. Tomasello, 1999; see Gopnik, 1999). The concept of shared intentionality, as well as other features that may construe a universal human socio-intentional motivation for music, is discussed in more detail in Chapter 4.

The culturally-enactive dimension - Merriam (1967) opens his ethnography of Flathead music with the statement that "all people, in no matter what culture, must be able to place their music firmly in the context of the totality of their beliefs, experiences, and activities, for without such ties, music cannot exist" (p.3). The veracity of this statement is widely accepted among ethnomusicologists. Even functional song types that are imbued unequivocally with biologically grounded affect, such as the lullaby (see above), exhibit considerable variation across cultures[27]. Although there appear to be some features of music that people can engage with across cultures, appropriate engagement is largely contingent upon cultural knowledge and immersion. Empirical evidence is unsurprisingly sparse given the difficulties inherent in such an approach. However, a few existing studies show that 'correct' engagement with pitch hierarchies (e.g. Castellano et al, 1984), and metrical structure (Stobart & Cross, 2000) for example, are dependent upon knowledge of culturally constructed structural schemas. Even individuals with considerable expertise in one cultural style can fail – in spite of tuition and a degree of cultural immersion – to perform seemingly simple tasks appropriately. Blacking's efforts to drum the beat as part of a Venda ceremonial musical performance, for example, were considered to be incorrect and even detrimental to the efficacy of the occasion (Blacking, 1973). Only after a considerable amount time spent with the Venda did he manage to engage more successfully. He notes, succinctly, that "to play Venda music you have to be a Venda" (p.92).

Crucially, cultural and associative processes go beyond conditioning an individual's ability to demonstrably engage with particular musical styles, but also relate to the meanings and affects attributed to music by individuals within a particular culture. An aspect of this is the 'darling-they're-playing-our-tune' phenomenon (Davies, 1978) which denotes the role of past experiences and co-occurring reference points in an individual's emotional reactions and attributed significance to musical songs (see Sloboda & Juslin, 2001). However, the term culturally-enactive implies much more than that and intends to encapsulate the fact that music's significance is not only affected by individual and social experience, but is also an active and generative component of culture. Cross clarifies that meaning and emotion in music are characteristics that "emerge from inter- and intra-personal processes that are shaped by – and that in turn may come to shape – the contingencies of their specific cultural matrix" (p.4). This mirrors equally crystallised quotations by Blacking (1969/73) who - with reference to

---

[25] To be clear cross-taxa correlations are *suggestive* of common biologically founded motivational-structural correlations. However, they are not an essential prerequisite in the identification of motivational-structural principles. Different evolutionary pressures may drive communicative repertoires in different directions and underlying biological 'codes' must be expected to be, in part, species-specific relating to, for example, particular anatomical features of the animal.

[26] Tylor's (1871), for example, defines human culture as "that complex whole which includes knowledge, belief, art, moral, law, custom, and any other capacities or habits acquired by man as a member of society". This remains a useful and operationally valid description and would appear to encapsulate what is being addressed here.

[27] Trehub and colleagues (1993b) show that, despite considerable cross-cultural similarity, North-American mothers tend to be much more blatant and demonstrative in terms of the affective content and structure of their lullabies. This correlates strongly with studies of ID speech which show that North-American mothers show greater intonational changes than mothers from Britain, France, Germany, Italy, Japan and China (Fernald & Marikawa, 1993; Papousek & Papousek, 1991) and has been seen as a reflection of North-American mothers' greater expressiveness and of the suggestion that some societies value soothing interactions more than arousing interactions (Toga et al., 1990).

the various individual and sub-culturally contingent roles and means of engaging (or not) with musical performance adopted by participants in Venda initiation ceremonies - notes that “music cannot make people act unless they are already socially and culturally disposed to act” (1969, p.39). In a later publication he explains that “rules of musical behaviour are not arbitrary cultural conventions and techniques of music are not like developments in technology” (1973, p. 99). Musical behaviour, he continues, “may reflect varying degrees of consciousness of social forces, and the structure and function of music may be related to basic human drives and to the biological need to maintain a balance among them” (Ibid, p.100). Effectively musical styles and their individual and collective significances are as varied and as complex as culture itself. The complex dynamic interplay between culture and music needs, in each instance, to be thoroughly understood in specific analysis and ethnography[28].

Applicability and relevance of the model - Perhaps the greatest strength of the dimensional model proposed is that it provides a clear and relevant reference point for broad biological and cultural comparative perspectives on music and other human and non-human communicative behaviours and repertoires. Broadly speaking we can expect the motivational-structural dimension to refer to communicative features/concepts that are applicable across a wide range of species. By contrast the socio-intentional can be interpreted as encapsulating human-specific yet universal/innately constrained facets of human communication. The culturally-enactive dimension in turn describes aspects of human communication whose meaning and affect result from generative and associative, culturally- and individually-specific processes and histories. A strong proviso, however, is that the model – as currently presented - doesn’t attempt to identify specifically musical features and could equally be applied to other human forms of communication. Whilst it is the generative whole and interactions between dimensions that potentiates a complete functional account of music, it is one of the central assertions of this thesis that an evolutionary perspective on music – with its requisite for describing human- and context-specific universals - needs to be focused on and around features that distinguish the specificites of a musical socio-intentionality. What is it that motivates us to structure the “communicative musicalities” of our dynamic mediations with others and the flow of our individual experience around the specificities of a musical pulse and/or musical tone? This critical question will be addressed in chapter 4.

Defining Music

It seems logical to conclude that if music is innately constrained, universally present and universally identifiable as “music” then it must, at least theoretically, be comprised of some universal features (or a family of resemblance) that identify and define it as such. Nevertheless, a clear and consensual definition of music has proved remarkably elusive and problematic. This, and even a reluctance to give operational definitions, remains a consistent problem in discussions relating to the role/significance of music and musical capabilities during the course of recent hominin evolution and, indeed, in the majority of cross-disciplinary investigations involving music. As we have seen from the above discussions music appears to be able to embody a multiplicity of functions and to have emerged as a combined result of a series of evolutionary adaptations. Hence it seems nonsensical to discuss the adaptive functionality of music or its status as an adaptation, exaptation or spandrel without first defining the matter for debate and its potentially unique features. It is these features (if any) that need to provide the primary basis for argumentation regarding the evolution and putatively adaptive qualities of musical capabilities. Otherwise, we are at constant risk of falsely attributing functionalities to ‘music’ that are in fact more precisely attributable to broader categories, such as, affective vocal communication. As discussed in the introduction the view, held by some researchers in the field (e.g. Hauser & McDermott, 2006; Mithen, 2006), that an actual verbalisable definition is unnecessary as we all know what we mean when we use the term ‘music’ is simply not tenable. Worse still, it operates as a barrier to clarity of understanding and the applicability of their own research as well as, if perpetuated, that of the field in general.

In general, authors writing broadly in the field of music psychology have been contented to give definitions of music to frame their particular article or argument. For example Stevens and Bryon (2016), discussing universals in music processing, define music as “temporally structured human activities, social and individual, in the production and perception of sound organized in patterns that convey non-linguistic meaning” (p. 15). This is, in many ways, a reasonable practical approach but one can easily imagine behaviours that would putatively qualify the definition but nevertheless would not be termed as music (e.g. mother-infant interactions). Alternately to frame their field of study ethnomusicologists have, for example, at times defined music simply as human communication outside of the scope of language. Bruno Nettl (2005) points out that such a wide net has enabled articles on such ‘non-musical” events as Koran reading, African drum signaling, whale and dolphin sounds, and Mexican Indian whistling speech to appear in the journal *Ethnomusicology*. Discussing many of the functionalities of music he offers a broad definition that is “intended to delineate those attributes that, in every community, appear to distinguish music from other spheres of human activity”. Readers are directed to the original articles for further specifics but he argues that something that might be identified as 'music' “appears to have some general attributes - roots in sound and movement, heterogeneity of meaning, a grounding in social interaction and a personalised significance, together with an apparent inefficacy” (p. 23).

A notable effort to define music *functionally* and in a way that highlights its significance to an evolutionary context, guards against ethnocentric conceptions of music, and enables its relationships to cultural and biological processes to be evaluated has been provided by Ian Cross (see below). He offers a definition that stresses three key features of music: embodiment,

---

[28] A particularly interesting and significant example is Svanibor’s description (1997) of the dynamic interface between music, context, identity, and culture in Croatia during recent turbulent political times. For a recent compilation of ethnomusicological approaches to musics in or as culture see Post, 2005.

entrainment, and transposable intentionality, which will be discussed separately below. This is, significantly, presented as an operational definition, without the objective of being to be either constitutive or essentialist.

Embodiment - Although this denotes broader conceptions of music being ‘physicalised’, the most overt demonstration that music is embodied comes from the fact that musical engagement, whether it is ‘active’ or ‘passive’, compels us to move. This is true to the extent that there is no dance without music (Nettl, 2005) and even to the extent that our cultural separation of dance and music can be seen to be an ethnocentric quirk (Arom, 1991). In all styles of music bodies produce music and in most (especially improvised styles) body actions are as much a determinant of the melody and rhythm produced as is any 'choice' of pitch patterns (e.g. Afgani dutars, Baily, 1985; blues guitar, Nelson, 2002). Equally choreography, whether it be directed toward sound production or not, is an essential component of performance and the creation of meaning (e.g. Japanese Taiko drumming). It appears that the connection between music and movement is so strong that even seemingly passive listening (cross-culturally, not the focal mode of engagement with music) activates areas of the brain associated with motor behaviour and sequencing (Janata & Grafton, 2003). This connection is, possibly, due to musical (in particular rhythmic) and timed and untimed motor behaviours sharing at least a subset of internal oscillatory mechanisms (Bispham, 2003; see chapter 2). In effect, music may be affective at a physiologically grounded level that is arguably not similarly accessible through other forms of communication.

Entrainment - In short, the process that allows us to be and to act together in time. In music, internal oscillatory mechanisms attune to regularities and periodicities in external stimuli providing reliable expectations as to the timing of future events and thus maintaining a co-dependent temporal relationship (e.g. Clayton et al., 2004). Entrainment has also been argued to provide a framework for interaction and expression in mother-infant and linguistic interactions (e.g. Malloch, 1999; Webb, 1972; Auer, 1999). However, degrees of regularities in music, and a constant coupling of action with perception, arguably endow the entraining phenomenon with a unique degree of permanence and unambiguity. While entrainment has been proposed as the root or enabling feature of human and animal temporal perception (Jones, 1976), sustained periodically structured synchronicity is rare in the animal kingdom (Merker, 2000). Additionally, it seems that we are the only species able to substantially and volitionally alter the temporal rate at which we synchronise (Repp, 2004). Similarly to embodiment, this form of action-perception coupled entrainment can be viewed as sharing overlapping internal oscillatory mechanisms with the motor system (Bispham, 2003). Through this connection and by providing a degree of interpersonal coupling (e.g. Benzon 2001), entrainment enables or least provides the foundations for a physiologically grounded group co-regulation of state. These issues, too, will be discussed in much greater detail in chapter 2.

Transposable Intentionality - Tolbert (2001, p.84) explains that 'music's power is not derived solely from syntactical or semantic referents, arousals and expectancies, or from its indexical relationships to a particular cultural context, but rather through its immediacy as a performance of socio-emotional essence and embodied gesture’. Effectively, and despite lacking direct linguistic referentiality, music would seem, nevertheless, to 'mean' something. Music can, for example, refer to communication with the spirits/dead (Feld, 1982), signal social transformation (Blacking, 1976), describe geographic associations (Will, 2002) or cultural narratives (Slobin, 1993). The key point that Cross is making by ascribing transposable intentionality to music is that meaning in music is both culturally constructed and individually flexible (depending on, for example, experience, social status or mood). In other words, it retains a high level of ambiguity and can mean different things to different people whilst remaining collectively meaningful (Cross, 2005). Cross quotes Goehr in suggesting that ambiguity is valuable within social and political contexts, as it ‘can be used to envision an alternative culture and political order while escaping the scrutiny of censor’ (Goehr, 1993). This resonates well with Turner's view (1974, 1991) that ritualistic or totemic acts occur most prevalently at moments when verbal description would highlight paradox and also explains music's cross-culturally prevalent role in ceremony and at times of political and social change. Coupled with its ability to create a co-regulation of states, music's transposable intentionality allows it to create feelings of togetherness and of shared experience in a way and in contexts that language, due largely to its referentiality, is unable to achieve (chapter 4).

In its most succinct form Cross (2003) proposes that *“music embodies, entrains, and transposably intentionalises time in sound and action.”* This definition is a welcome addition to the literature and certainly appears to broadly encapsulate biological and cultural perspectives and to be applicable to - and descriptive of crucial components of - all known musics. However, it does not, perhaps, clarify sufficiently the specificities of music and fails, in the manner it is presented, to account for specifically “musical” means of engaging with pitch. None of the three constitute parts are in and of themselves peculiar to music. Language, poetry and ritual are all, to varying degrees, flexible to interpretation of meaning; entrainment is, broadly defined, inherent in a range of interactions and coordinated group behaviours; and embodiment is also a necessarily broad concept. As such, it could perhaps be interpreted as a substantive statement rather than as a definition per se. The statement offered by Cross effectively serves a very different function than a resultant narrower definition focused on the central aim of this thesis to identify and describe the constitent parts of the narrow faculty for music – those psychological, physiological, behavioural, and/or cultural features that are both universally present and yet specific to music. This thesis aims therefore to build on this proposed statement by adding an increased degree of specificity and aims to contribute towards addressing and extending those partially definitional questions I posed in the title to an article on rhythm (Bispham, 2006): What is Music? Who has it? And Why?

Debates and hypotheses on music and evolution need to be focused on, grounded in, and informed by, an understanding of the specificities of musical interaction and engagement – the narrow faculty for music. Otherwise, we are at constant risk of falsely attributing functionalities to “music” that are in fact more precisely attributable to broader universal categories of relevance.

SECTION II – The Narrow Faculty for Music - Musical Motivation, Musical Pulse, and Musical Tone

INTRODUCTION

Hockett's "Design Features"

A complex communicative form such as language or music is, necessarily, an intricate mosaic of capacities. Individual traits may have influenced evolutionary fitness for direct adaptive consequences of that particular complex behaviour or not. Nevertheless at some point in time capacities emerged according to and in response to evolutionary principles and pressures. An important background to the methodological approach in this thesis comes, therefore, from the literature on the evolution of communication, in particular the influential contributions of Charles Hockett on language. Hockett proposed a "design-feature"[29] approach to comparative linguistics in which he endeavoured to identify the similarities and differences among animal communication systems and human language. In a paper published in 1959 - Animal 'Languages' and Human Language - he initially proposed seven features. These were gradually augmented until, in perhaps his most well-known exposition of this approach - "The Origin of Speech" (1962) - he settled on 13 design-features: Vocal-Auditory Channel; Broadcast Transmission and Directional Reception; Rapid Fading (transitoriness); Interchangeability; Total Feedback; Specialization; Semanticity; Arbitrariness; Discreteness; Displacement; Productivity; Traditional Transmission; Duality of Patterning[30]. He argued that only human spoken language utilized all 13 concurrently but that all human and animal communication systems share some of the 13 design features. He further posited that these could differentiate human spoken language from animal communication and other human communication systems such as written language. Traditional transmission and duality of patterning, in particular, were regarded as being key to human language. Hockett's system was highly influential in comparative studies of communication. For example, many ethologists in the 1970s applied the design-feature approach to the study of a range of animal communication systems (e.g. Marler 1970; Thorpe 1972; Hinde 1975). Hockett's classification system also appeared frequently in broad semiotic and biosemiotic literatures as a foundational attempt to systematically differentiate between human and non-human communication (see e.g. Danesi and Perron 1999: 109–111; Martinelli 2010: 221; Nöth 1990: 155–156). As such, although this thesis is pursuing a different goal of specifying features that are specific to music (in essence unique "design features"), it is fundamental to my argument that language and music can be described as a conglomerate of "design features" that can individually be subject to comparative analysis. Notably, there have been some dissenting voices. Everett (2005), for example, presents a detailed linguistic analysis of the Pirahã language and argues that this challenges a direct application of Hockett's design features by showing that interchangeability, displacement, and productivity may be culturally constructed and constrained. Also of particular relevance, in terms of similar methodological concerns regarding music, is Wacewicz & Żywiczyński's assessment (2015; Wacewicz 2012) that highlights and critiques "a focus on the means at the expense of content and an emphasis on the code itself rather than the cognitive abilities of its users" (p. 29). They propose that

> "the field of language evolution requires and presupposes a more "cladistic" approach to language: as a suite of sensorimotor, cognitive and social abilities that enable the use but also acquisition of language by biological creatures" and note that "most of those in the field see language as a complex (or mosaic—Hurford 2003) of cognitive skills, or an even more multifaceted phenomenon, grounded in but transcending individual cognition (e.g. Gärdenfors 2004)" (Ibid).

The Human Faculty for Language - Methodological Concerns

These preceding points are addressed in a more recent, and highly influential, paper addressing broad methodological concerns for an evolutionary perspective on the faculty of language. Hauser, Chomsky and Fitch (2002) suggest a distinction between what they term a faculty of language in the broad sense (FLB) and a faculty of language in the narrow sense (FLN). Essentially FLN seeks to describe capacities that are unique to humans within the context of language. As such, it is a direct fit, inspiration, and methodological analogy to how a narrow faculty for music is framed in this thesis (see below). FLB includes an internal computational system (FLN, see below) combined with at least two other organism-internal systems; 'sensory-motor' and 'conceptual-intentional'. The sensory-motor system is posited to include skills such as vocal imitation and invention; neurophysiological systems of action-perception; discrimination language-relevant sound patterns; potentialities of vocal tract anatomy; and the modalities of language production and perception. The conceptual-intentional system in turn incorporates capacities to attribute others' mental states ('theory of mind'); the capacity to acquire nonlinguistic conceptual representations; referential vocal signals; intentional and rational imitation; and intentional voluntary control over signal production. Accepting that opinion is divided on the precise nature of these systems, and whether they are substantially shared with other vertebrates or uniquely adapted to the exigencies of language, the authors take as uncontroversial 'the existence of some biological capacity of humans that allows us (and not, for example, chimpanzees) to readily master any human language without explicit instruction'. FLB includes this capacity but excludes organism-internal systems that are necessary but not sufficient for language such as memory, respiration, digestion, and circulation. Leaving the door open to more inclusive definitions[31] the authors hypothesise that FLN includes a

---

[29] A dictionary definition of design features in an evolutionary context may describe, for example, "an organismal trait that can influence rates of death and reproduction, and hence Darwinian fitness" (McGraw-Hill, 2003).

[30] In conjunction with anthropologist Stuart Altmann 3 more were added in a report presented in 1968:
Prevarication: A speaker can say falsehoods, lies, and meaningless statements.
Reflexiveness: Language can be used communicate about the very system it is, and language can discuss language
Learnability: A speaker of a language can learn another language

[31] Lieberman (1996), for example, has argued that the sensory-motor systems were specifically adapted for language and thus should be a part of FLN. Recursive structure vs process – so in practice not really recursion ad infinitum but perhaps an extended form of working memory (see Read 2008)

computational system (narrow syntax) that generates internal representations and maps them into the sensory-motor interface by means of the phonological system and into the conceptual-intentional interface by means of the (formal) semantic system. They suggest that the crucial component of FLN is recursion, which they identify as a computational system operating primarily over the syntactical domain. FLN thus takes a finite set of elements and yields a potentially infinite array of discrete expressions[32]. According to the authors this core recursive aspect of FLN appears to lack any analog in animal communication and possibly other domains as well.

Hauser, Chomsky and Fitch strongly favour a comparative approach to studying the evolution of the faculty of language and promote the quest to explore both direct evolutionary homologies resulting from shared ancestry and analogies occurring as traits emerge in separate lineages as similarly resultant functions in response to comparable selective pressures. They propose, on the basis of a wide range of comparative empirical data that only FLN is uniquely human and hypothesise that most, if not all, peripheral components of FLB are based upon mechanisms shared with other animals, with differences of magnitude or quality rather than kind. On the other hand, FLN - the computational mechanism for recursion - is viewed as being recently evolved and unique to humans. The authors explain that

> "according to this hypothesis much of the complexity manifested in language derives from the complexity in the peripheral components of FLB, especially those underlying the sensory-motor and conceptual intentional interfaces, combined with sociocultural and communicative contingencies. FLB as a whole has an ancient evolutionary, long predating the emergence of language and a comparative analysis is necessary to understand this complex system. By contrast, FLN as a uniquely human capacity must have evolved in the last 6 million years since our divergence from the last common relative of chimpanzees and humans." (p. 1570)

Importantly, the authors recognise that both FLB and FLN may have evolved for plausible communicative precursors or for reasons outside of linguistic/communicative contexts such as numbers, navigation and social relations. They propose therefore that comparative studies might look for evidence of such computations outside of the domain of communication[33]. Furthermore they view shared mechanisms across different cognitive domains (in particular music) as a valuable source of information about the structure of the faculty of language. An important point to make is that they distance this suggestion from debates and potential criticisms of the so-called minimalist program in linguistics, preferring in this instance be remain largely non-committal as regards the virtues and faults of the various flavours of generative grammar currently available. As they later clarify "the only assumption made in HCF [Hauser, Chomsky, & Fitch] and here, about syntactic theory is the uncontroversial one that, minimally, it should have a place for recursion" (Fitch et al., 2005, p. 183).

In this later paper (Ibid), in response to some criticisms from Jackendorff and Pinker (2004), they also emphasise the central impetus for the original article in a quotation that is directly pertinent and, hence, worth quoting here in full:

> The primary motivation for writing HCF was the recognition of some pervasive confusions that have led to persistent and unnecessary misunderstandings among researchers interested in the biology and evolution of language. Such misunderstanding has polarized debate unnecessarily, has helped to fuel dogmatic and even hostile stances, and has generally acted to block progress in this field, including especially the severing of possible collaborative projects between linguists, psychologists and biologists. It has contributed to a situation in which animal researchers interested in language almost automatically consider themselves anti-linguist, or anti-generative, while some linguists feel justified in being anti-cognitive or anti-evolutionary. The FLN/FLB distinction, we hoped, would help the field to see that there is no incompatibility between the hypotheses that FLB is an adaptation that shares much with animals, and that the mechanism(s) underlying FLN might be quite unique (p.193).

It is this motivation as well as the proposed methodological need to fractionate language into its component mechanisms; the need for an empirical approach to test individual hypotheses about these mechanisms; the value of comparative data from

---

[32] Recursion is perhaps not the most precise term to use as it implies infinity. However, we (as the authors recognise) are restricted by other organism internal factors such as lung capacity, working memory, concept formation and motor output speed and hence exhibit a limited capacity for recursion. Effectively they are describing a process rather than a recursive structure which may in practice reflect an extended form of working memory (see Read 2008).

[33] Implicit, although not explicitly mentioned, in this approach is the notion that linguistic ability requires processing information across different cognitive domains as well as across modalities. In other words for a faculty of language to operate it requires not only the individual components but also some form of integrative fluidity. Mithen (1996) has suggested, following Fodor's distinction between innate, specialised cognitive competencies and a general, central, intelligence, that the course of human cognitive evolution is *uniquely* characterised by an increase in the capacity to integrate information across domains and to transfer information or skills between domains. Mithen, with the support of the archaeological record, describes three evolutionary phases:

Phase 1: Minds dominated by a domain of general intelligence - a suite of general-purpose learning and decision-making rules.
Phase 2: Minds in which general intelligence has been supplemented by multiple specialized intelligences, each devoted to a specific domain or behaviour, and each working in isolation from the others.
Phase 3: Minds in which the multiple specialized intelligences appear to be working together, with a flow of knowledge and ideas between behavioural domains (p.64).

The flexibility of phase 3, which according to Mithen characterises the modern mind, is seen to have emerged only within the last 100,000 years and becomes most evident in the archaeological record with the "cultural explosion" that occurred some 30,000 to 60,000 years ago. Although music is not explicitly mentioned by the author in the original exposition, it has particular relevance to music as musical ability also requires processing information across domains and modalities (see Cross, 1999).

diverse animal species; and the need for collaborative, inter-disciplinary work in this endeavour that inspires the methodological approach taken in this thesis. Broadly speaking the same concerns are shared in the study of music in evolution. The particular theoretical conclusions and detailed discussions on the evolution of language are beyond the scope of this thesis. Nevertheless, by way of ingesting some further interpretational guidance and understanding some potential analogous theoretical complications and pitfalls it is worth very briefly considering the proposal that FLN is characterized solely by recursion and is the only uniquely human component of the faculty of language.

Interpreting the evidence – theoretical concerns

A first challenge is that although the authors describe FLN as being "composed of those components of the overall faculty of language (FLB) that are both unique to humans and unique to or clearly specialized for language", (2005, p. 182) it is debatable whether recursion can be considered as having arisen as an adaptation for language, music perhaps, or something else. In fairness the authors made this point clear stating also that "something about the faculty of language must be unique in order to explain the differences between humans and other animals—*if only the particular combination of mechanisms in FLB*" (2005, p. 182) (emphasis added). They also note that the contents of FLN are as yet to be empirically determined. Critically, it is not just a question of curiosity whether the contents of FLN are shared with other domains. Eventually proposals on this matter and subsequent debates inevitably collide with fundamental philosophical perspectives of what language is and ultimately to the core of who we are. Is language *merely* a form of communication - an extension of prior systems - or is it in essence a fresh mode of being and consciousness? Is it, for example, the by-product of, or even an aspect of, the evolution of symbolic representation (e.g. Deacon, 1997)? Or is it even primarily a mode of thought – an internal language with its uniqueness captured in the uniqueness of its generative nature of thought, its universal grammar (Chomsky, 2010). Is it instead, as Daniel Dor (2015) suggests, definable as a 'sharing of experience', as the communication of thought with its distinction deriving from the nature of the ideas we wish to communicate? Alternately again, could languages be best conceived as tools to solve the twin problems of communication and social cohesion? – "Tools shaped by the distinctive pressures of their cultural niches - pressures that include cultural values and history and which account in many cases for the similarities and differences between languages" (Everett 2012, p. 6).

Even if we were to accept the proposal that recursion is the core component of FLN it would still leave many of these centrally reflexive questions unresolved, especially those regarding the sequence, niches, and pressures of evolutionary adaptation. Corballis (2014) argues, for example, that it was the emergence of the capacity for recursion that first distinguished us in the animal kingdom and provided the foundation for subsequently defining characteristics. In his argument it is recursion that enables us to conceive of our own minds and the minds of others and which gives us the power of mental "time travel". It affords, in his view, the ability to insert past experiences, or imagined future ones, into present consciousness. In turn he posits that recursive structures led to the emergence of language and speech, which ultimately enabled us to share our thoughts, plan with others, and reshape our environment to better reflect our creative imaginations. To highlight further potential complications, even in this argument we can potentially be left with questions as to whether this represented an initial qualitative shift from working memory capacities in nonhuman species (see Read 2008; Wynn and Coolidge, 2007) or a more abrupt restructuring of central cognition.

Many of these questions are considerably beyond the scope of enquiry in this thesis. However, as well as drawing inspiration from the methodological advocacies, it is interesting to consider the scope of interpretational debates, concerns, and philosophical challenges that have been advanced in studies of the evolution of language. These have interesting parallels and comparable complexities in music and will be important to consider carefully in future extended discussion on proposals made for a faculty for music in the narrow sense (FMN). In many regards this section (Chapters 2-4) - constituting the central claims of the current thesis - will follow analogously the suggestions on how to describe and comparatively support a narrow faculty for language (FLN). However, there are also some critical differences in our quest to search for the specificities of music. These are due in part to inherent differences between music and language and in part to the current state of enquiry in the respective fields. In posing the question 'what (if anything) defines and distinguishes musical engagement from other forms of non-human and human behaviour?' we are addressing and probing a potentially even narrower conclusion.

Music and Language – a Communicative Continuum

In general it makes perfect sense to follow a similar comparative method as that suggested for language. Music and Language putatively share a common evolutionary past until a relatively recent divergence. Both are complex, uniquely human capacities made up of a vast array of largely overlapping subskills. The commonalities between the two domains are particularly abundant when we assent to a broad conceptualization of language in which the exchange or conveyance of semantically decomposable 'information' is merely a part of the whole. It is, as music also is, primarily a multi-modal form of structured interaction. Language and music both incorporate intrinsic motive pulses, gestures, affective expressions and regulations, subtle cues of body language, shared goals and intersubjectivities. Both are combinatorial, generative, relational and concerned with the maintenance of social relations. Both are essentially manifestations of the human capacity to communicate. In fact a select group of influential researchers (Cross et al, 2013) has proposed that the two constitute a continuum rather than discrete domains. This continuum is proposed to be interpreted in terms of at least two dimensions, the first running from definite to indefinite meanings and the second from greater to lesser affective potency. They give the example of the Suya people of the Amazon (Seeger, 1987) to highlight that, although we in modern Western societies categorise music and language as discrete and principal communicative genres, other complex communicative taxonomies in which, for example, social function and context are the primary distinguishing characteristics are possible and operational in traditional societies (see also Basso 1985; Feld 1982; Lewis, 2013)[34].

Cross et al (2013), following List (1963), argue within our contemporary Western taxonomical system that "overall, no single criterial attribute, save perhaps that of propositionality, distinguishes between language and music clearly and comprehensively. Language, they continue, "unlike music, provides a way of sharing information about states of affairs by means of truth-conditional propositions and thus of coordinating action. It enables mapping between worlds, thoughts, and selves, the formulation and exchange of information, and the coordination of joint, goal-directed action" (p.545). They do also cautiously highlight that "the discrete tones and pitch sets that supply grist for the musical mill in most cultures are rather unique to music" and that "there are differences in the regularity of timing between most registers of speech and most genres of music, with the latter featuring *explicit* use of isochrony" (Ibid, emphasis added)[35]. However, pitch, pitch contour, "rhythmic" regularity, temporal structure and emotional expression and regulation are, broadly speaking, also critical features in oratory, poetry, linguistic and parent-infant interactions and also in certain animal communications. Therefore, to definitely isolate music on the basis of any of these attributes we will need to be much more specific. Essentially, as well as analogous inspirations for research into the faculty for language, an additional motivation for describing a narrow faculty for music in this thesis is also to explore if it is possible to pinpoint features that positively distinguish music from language and other communicative repertoires. The alternative to addressing this impetus is that arguments regarding the role of music in human evolution, its modern-day functions and particular efficiencies, and our appreciation of its position within broader fields of human cognition and behaviour will remain, necessarily, somewhat vague and lacking in clarity and consensual understanding.

Appoaching a Narrow Faculty for Music

Seemingly not tied down by one central functionality, music is substantially more flexibly manifest than language and also more adaptably intentional and significant. The wonderful variety of musics across the globe defies observations of common ground even as we instinctively sense "musicness" and adherence to our taxonomic enclosures. Recent avant-garde experimental genres such as the *musique concrète* of the 1940s are among the many examples of how almost any sound and/or action can potentially be abstracted, shaped and transmuted into being something we regard as music. It is a central supposition of the argument of this thesis that the potential psychophysiological "contents" given space in musical form are immeasurably mutable. In practice, of course, certain instantiations may lend themselves more readily to our socio-intentional drives than others. As Storr (1992) notes, although musics are indeed hugely diverse, they are less so than one can imagine organising sounds and movements. Nevertheless, given the scope of potential and observed musical form within and across cultures it seems that any description of commonality needs to be addressed at the level of psychological mechanisms that anchor a musical form or structure. As discussed in the introduction and in chapter 1, there are cogent, logical and bio-cultural reasons to assume the existence of distinguishing features at the very least at the level of implicit categorisation by musicologists. This argument persists even if, as seems likely, we are commonly only aware empathically of a conjoining experience. This latter point in particular highlights the challenge inherent in the goals of this thesis. Fortunately, however, combining all three comparative perspectives that are integral to the question of the constituent parts of a narrow faculty for music – seeking cross-cultural universalities, cross-species homologies and analogies, and intra-species domain-specificities - rather than exponentially complicating the issue, appear to converge on similar suggestions thus adding weight to the conclusions reached.

Another supposition posited throughout this thesis is that these universally defining features of music are shared psychological mechanisms that structure attention and motivation and are not always fully physicalized or observable (see introduction). This approach can be envisaged, for example in how we understand findings such as the recently titled 'speech to song illusion'. In this paradigm repeated speech patterns start to be classified as being music by experimental participants. This has been interpreted as physical evidence for a somewhat nebulous divide between 'musical' and 'linguistic' rhythm perception. I argue, however, that this reflects a relatively clear transition, due largely to increasing predictability, from one form of processing to another and an identifiable psychologically grounded specificity of musical rhythm. Music ultimately happens in the mind. It is formed and understood based on the psychological and attentional mechanisms which structure our thoughts, actions and interactions in time. The approach taken in chapters 2-4 is therefore based primarily upon empirically grounded psychological studies of musical capabilities and engagement. The physical products such as the sounds produced by isochronous tapping or sustained humming are described, in turn, as acoustic correlates of engagement. This has the additional benefit that perception and production - which anyway are in essence inseparable in music (and dance) - are mutually encapsulated in the term 'engagement'.

---

[34] The Suya describe three form of communication: Kaperni most closely resembles interactional "speech" with text the predominant (but variably so) priority'; Saren is "instructional" and more formalized declamatory or empowered human speech with a priority of relatively fixed texts over relatively fixed melodies; and Ngere, which we might regard as "song", in which melody is fore fronted and in which time, text, and melody are set and of nonhuman liminally powerful origin ('authorised' for example by current power structures or past invocations). Importantly these forms can leak into each other with occasionally loose demarcations between the forms. Seeger 1989 also notes the different concept of *Ngere* to our understanding of "song" stating that "*ngere* means both a song (melody) and the movements that accompany it. Stance and dance are thus an integral part of music, all part of a single communicative act called *ngere*. At a more specific level of contrast *ngere* means a unison song as opposed to the *akia*: at a more general level *ngere* means "ceremony", including the whole range of activities, songs and movements that characterize Suya musical occasions (p. 375).

[35] Isochrony describes a steady, almost metronomic, beat

I will also attempt to describe the psychological devices involved in the absolute minimal organisational and motivational characteristics for something to be "musical" - root capabilities involved in engaging with a simple musical pulse, a musical tone and the most intrinsic motivations to do so. Without wishing to artificially restrict the scope of discussion these ostensibly most "simple" forms of pulse, tone, and motivation can perhaps be characterised as the capabilities needed to cooperatively tap an isochronous pattern, to interactively hum a matched or harmonised sustained pitch, and an inherent desire/impulse to do so. It is hoped also that by focusing on such "basic" foundational capabilities we are analysing, as much as possible, at levels that reflect biological and ontological constraints and which operate to support cultural creativity rather than reflecting specific acculturations and learning. Nevertheless, a potential caveat to note from the outset is that some cultural influences may nevertheless be inherent and that the overwhelming majority of experimental work in this area has been undertaken with Western subjects. Acculturation has variously been shown to affect even the simplest musical tasks. Arguments have even been presented for some prenatal influence (e.g. Panksepp, 2009). Hence some caution is required to avoid ethnocentric interpretations and suggestions. Such attentiveness is inherent in the further course of this section which, giving particular weight to (the relatively few) available cross-cultural studies, aims to offer cross-culturally generalisable concepts and hypotheses.

It is absolutely critical to be clear, however, that this is not meant to represent a reductionist or minimalist view. These root capabilities are not considered to be quasi-grammatically generative nor functional in isolation. They are instead merely a psychological space, an individual or interactional framework in which a full range of human vitality and experience can be shaped, abstracted, transformed in time and given a particular kind of meaning and significance. Furthermore, as we shall see, they are manifestly embedded in broader capabilities and the complex mosaic of our evolutionary past. Even these seemingly basic skills are intrinsically embedded in and operationally dependent upon capacities for movement, vocalization, intersubjectivity, sociality (e.g. Himberg 2015) and culture, as well as direct and resultant affectivities. Ultimately, of course, it must be these connections that ground any functional account for music and its particular efficacies. Musicality, to quote Colwyn Trevarthen (1999/2000) "cannot be synthesised or accreted in the mind from separately anarchic processes, but must have one coherent and orderly rhythmic foundation" (p. 160). One can argue more specifically, for example, that all of the world's music may have ultimately arisen out of some combination of vocal utterance and coordinated action (Clayton, 2009). Music in its full forms ultimately then "communicates the mind's essential *coherence of purpose*, a willfulness that holds its elements in a narrative form through phenomenal, experienced, time" (Malloch & Trevarthen 2009, p. 1).

The argument is set out and structured over the next three chapters on Musical Pulse; Musical Tone; and Musical Motivation respectively. A temporal organization of action (i.e. pulse and rhythm) and frequency-based order (i.e. melody and harmony,) are widely considered to be fundamental, descriptively and neurophysiologically distinct (see Peretz and Zatorre, 2005), features of music. Few, I think would argue, therefore, against an initial division into pulse and pitch as likely sources for music's unique characteristics. Brown's musilanguage model, for example, also argues for pitch blend and metric rhythms as being "design features" of music that "reflect music's origins as group selected trait" (p. 4; see chapter 1; Conclusion). To fully encapsulate an inherent capacity observable in *all* musics, specific features of musicality are described as a species- and context-specific ability to engage with *configurations* of musical pulse and/or musical tone. This incorporates the fact that not all musics employ an observable temporal pulse (Clayton, 1996); that some musics are entirely non-pitch based; and that an engagement with musical pitch need not imply melody, harmony or the use of scales (e.g. monotone chanting). The concept of motivation, although rarely discussed in musicological literatures, must also be considered a crucial component of human musical engagement. In the broadest and most inescapable sense all human (and non-human) actions are necessarily motivated by the pursuit of certain "goals". In other words we are constantly motivated towards maintaining or altering our physical, cognitive and/or affective environment – our experiential consciousness. These three features are presumed to provide symbiotically a more or less invariant framework for individual experience and/or interpersonal interaction – a coordinating strategy[36]. In short the claim made is that music is universally and context-specifically built upon and identifiable by the enacted fundamental capabilities and acoustic correlates of engaging with musical motivation and configurations of musical pulse and musical tone.

Drawing upon psychological and neurophysiological literature, I have previously suggested that musical pulse and musical tone be viewed as psychophysiological mosaics - constellations of concurrently operating, hierarchically organized subskills (Bispham, 2003; 2009). I further posited that these subskills share overlapping internal oscillatory mechanisms (Bispham, 2003) and that, as such, all subskills, as well as the resulting constellation, can be viewed as being grounded in, and as having exaptively evolved from, fundamental kinesthetic abilities and modes of perceiving temporally organized events. This may, perhaps have presented as an overly disjointed account, not fully reflecting the embeddedness and sometimes nebulous divides between subskills. I should like to clarify, therefore, that these components are not intended to be understood as being fully autonomous. Broadly speaking pulse and tone are integral components in all communicative repertoires so the question becomes one of whether there are aspects that are distinct in a musical pulse and/or tone. Viewing musical pulse and musical tone in this way provides an avenue for exploring useful comparisons with other types of human and animal behaviour as well as critically pointing toward putative evolutionary continuities. This approach mirrors the proposals for the study of the evolution of language discussed above (Hockett, 1962; Hauser et al., 2002), suits ideas of increased cognitive domain-flexibility in the hominin lineage (Mithen 1996), and fits perfectly the predominant view in evolutionary theory that complex behaviours such as music evolved in a mosaic fashion, with individual components emerging or evolving independently or for independent reasons at times, and/or reforming with other components at other times (Foley, 2005; 2016). The following cross-species and intraspecies comparative perspectives, by focusing on

---

[36] The term 'strategy' is used to incorporate the important detail that an engagement of musicality is partially volitionally controlled.

individual behavioral components, are implicitly set within this constellation framework. In the following chapters putative subskills will be subject to comparative analysis and discussion with a particular view to identifying those features that are specific to music.

Extending from previous publications (Bispham 2003; 2006; 2009; Cross et al., unpublished) musical pulse is described comparatively in conjunction with discussion on general timing abilities; the perception of time; absolute vs. relative temporal processing; periodic and nonperiodic action; a coupling of action and perception; sensorimotor synchronization; and error correction mechanisms. In particular I will highlight the challenges of relative processing in nonhuman primates and the notion that reported instances of analogous pulse-based behaviour in non-human species need to be challenged with reference to our understanding of the psychology of sensorimotor synchronization in humans. Perspectives that posit forms of structural and motive pulses in non-musical human interactions are also anaylsed accordingly. Period correction mechanisms, a corresponding awareness of a pulsed framework, and the afforded sustained nature of attention will be considered to distinguish musical pulse from temporal structuring in other communicative domains.

In similar fashion to the preceding section on pulse, musical tone will be correspondingly deconstructed into capabilities of controlled vocalization; pitch perception; relative pitch processing; psychophysiological correction mechanisms; frequency attunement; and harmonic tensions. Structural concepts such as scales, tonal hierarchies, and harmonic systems are described as parsimonious cultural constructs based upon these “root” capabilities of musical tone. Analogously to the conclusions on musical pulse, a desire to achieve and maintain certain frequency relationships and a resultant sustained stable pitch centre is seen to provide a mutually manifest framework for interaction. It will also be argued similarly that correction mechanisms based on the output of self or others (and again perhaps particularly those that invoke an awareness of the structural framework) seem to be critical in distinguishing musical tone from uses of pitch in other modes of human and animal communication.

The effort to describe and distinctively identify musical motivation necessarily follows a somewhat more discursive structure starting with ethnomusicological perspectives music’s multifarious uses and functions. The seemingly most likely intrinsic motivations that find voice in music are discussed and compared to other behavioural and communicative channels. The literatures on emotional expression and the individual and interactional regulation of affect in music are reviewed and critiqued. Although extended and often intensified, these subjects are not viewed as being intrinsically special to music. Following, Juslin and Laukka (2003), comparative analysis is argued to suggest that music shares common acoustic and structural cues with emotional vocal communication. A potentially specific or pronounced efficacy of music is posited in the enactive regulation of mood. Interactional intersubjectivities are also explored in terms of whether music provides some unique access to the mental states of others and to what extent we share in each other’s experiences and goals in musical performance. The broad psychosocial implications of synchrony, entrainment, group action, and ritual are similarly examined. I will suggest that the specificity of musical motivation lies in endeavouring towards a sustained attentional structure - and extended phenomenological present – that affords an efficient regulation of mood, coordinated group interaction and a consequent drive towards a convergence of motivational states.

Following efforts to identify features that represent a uniquely emergent capacity for music I will argue that the constituent parts of a narrow faculty for music – those psychological, physiological, behavioural, and/or cultural features that are both universally present and yet specific to music – provide a stable temporal framework for the efficient management of individual behaviour, dyadic and/or group interaction and are primarily concerned with sustaining attention and extending the experience of the perceptual present.

Chapter 2 - Musical Pulse

I have previously written similarly motivated comparative accounts on 'interpersonal musical entrainment' (Bispham, 2003), 'musical rhythmic behaviours' (Bispham 2006a; 2006b) and 'musical pulse' (Cross et al, 2008; Bispham 2009). Since the publication of these papers the issue of pulse and entrainment in music has attracted a huge amount of interest across a wide range of academic disciplines and has been the subject of numerous studies and publications. In particular some cleverly designed infant/child developmental studies (e.g. Philips-Silver & Trainor 2005; 2007); an increasing number of ethnomusicological accounts of entrainment (e.g. Clayton 2009); neurological hypotheses of neural resonance (Large et al, 2015); studies of neural networks in human and non-human beat perception and synchronization (Merchant et al, 2015); and a range of experimental investigations highlighting how even simple instantiations of pulse-based entrainment are functionally embedded in our central prosocial drives and motivations (e.g. Kirschner & Tomasello 2010) have vastly expanded and profited the literature in this area. In parallel, research into related matters such as synchronous action (e.g. Gill 2016), cooperation and empathy (e.g. Rabinowitch, 2015), and physical gesture (e.g. Kendon, 2004) have also broadened the scope of functional interpretation regarding forms of interactional entrainment. An increasing number of examples of animal 'rhythmic behaviours' have also been subject to much analysis and debate (see Fitch 2015; Wilson and Cook, 2016). Most notable of these has been 'Snowball' the dancing (headbanging) cockatoo who has fascinated researchers, divided opinion and even caught the public imagination via youtube and appearances on popular television shows.

Some authors have argued that the evidence has mounted conclusively against some of my previous suggestions. Rouse et al., 2016) for example state the "phenomenon of "beat keeping" was believed to be unique to humans (Wallin et al., 2000; Bispham, 2006; Zatorre et al., 2007), but new findings in non-human animals have decisively put that idea to rest" (p. 257). However, this misrepresents my conclusions and analytical approach. First of all, in terms of the capacity for entrainment, I supported evolutionary continuities from more basic temporal and motor functions arguing that "temporal processing, periodical behaviour, error correction mechanisms, dynamic attending and motor skill/coordination are all subserved by a common internal periodic oscillatory system". Furthermore I described that "IME (interpersonal musical entrainment) can operate at temporal levels below perceptual thresholds suggesting that the internal processes which control the timing of synchronised action are directly coupled to the sensory information without any mediation by awareness or perceptual judgement" (Bispham, 2003 p. 34.) Evidence and support for this suggestion have gained considerably in over the last years and the fundamentally rhythmic nature of motor and sensory processes will again be discussed here. Secondly, I did not propose anything as absolute as claiming that no non-human animals had any beat-keeping abilities. Instead I was aiming to isolate features of a "musical" pulse that distinguished it from the many other forms of pulse and rhythmic structurings that are inherent in a range of human and non-human behaviours and communicative forms (Bispham, 2006). A methodological concern that remains pertinent was that, in order to identify a definite homology or analogy in animal behaviours, we need to confirm similar psychological mechanisms as well as a similarly observable physical product. The conclusions I put forward focused particularly on period correction mechanisms and a corresponding awareness of the temporal framework and the sustained nature of engagement. I argue that, although the relevant literatures have advanced considerably, these suggestions still hold weight and will again feature here, as will the notion that musical pulse is primarily rooted in providing a temporal framework; collective emotionality; a feeling of shared experience; and cohesiveness to group activities and ritualistic ceremonies. However, I now also propose that this perspective needs to be broadened further to include both the management of interpersonal interaction *and* individual behavioural and experience.

The first few sections of this chapter will briefly describe our central experience of time and timing mechanisms. I will highlight the notion that a musical pulse is built upon these core mechanisms, but is not distinct at these fundamental levels of cognition and experience. Of particular importance is the evidence presented for a time span that commonly corresponds to the perception of a perceptual present. Later argument in this chapter will show how a musical pulse and corresponding mechanisms for period correction afford a constantly renewing link from one perceptual present to another. The subsequent section on dynamic attending theory is also critical to understanding the mechanisms involved in a musical pulse and why it is, in particular, an attentional structure. This theory describes processes that are inherent in any attentional pulse. Hence, these also are not, I propose, intrinsically unique to a musical pulse. It is, instead, the particular generative features of attentional structure in a musical pulse that will be argued to be distinct later in the chapter.

The Experience of Time – Temporal Ranges

It is indicative of what music is, how it is embedded in the central features of our being and experience, that in trying to describe the basic building blocks of a "simple" musical pulse we are, from the outset, reliant upon describing mechanisms that form our essential understanding of the flow of life events and phenomenological experience. Reflections on time and experience are, of course, long-standing and highly complex. Over 2000 years ago Aristotle deduced the continuity of time – its infinite divisibility – from the continuity of motion. This subsequently was deduced from the continuity of the space negotiated. Time here is, as Gale describes, "made continuous by the indivisible, present-now moment, which links the past to the future by serving as the termination of the past and the beginning of the future" (2016, p.1). This is only a starting point to intense philosophical debates, straddling our central understandings of physics and psychology, which are far beyond our current scope of enquiry. However, we can understand this description of continuities being experienced in conjunction with our fundamental nature to perceive temporal ranges as being simultaneous, sequential, flowing/happening, present, experienced or anticipated. Pöppel (1999, 1999/2000), for example, describes three basic temporal experiences:

simultaneity; non-simultaneity; and temporal order; and has suggested that the temporal processing of sequential information can be classified into four temporal ranges[37].

London (2012) offers a comprehensive discussion on the relevance of these temporal ranges in music. The most crucial for understanding the structure and phenomenological experience of a musical pulse corresponds with a range categorised (e.g. Block and Gruber, 2014) as a time span of roughly **30ms-3-5secs.** This is both the time range within which we can feel and follow a "beat" in music and in which, more broadly speaking, events appear as a changing present – there is an experience of time flowing and experience happening. This has been supported, for example, in paradigms of film-frames/snapshots being presented at increasing interstimulus intervals (e.g. men walking and bread being toasted [Gruber & Block 2013]). Participants generally reported that the flow of time starts to get lost at about 3secs and has completely disappeared by 7secs. The same authors also presented subjects with the first four notes of Beethoven's fifth symphony and reported that people no longer recognised it at an interstimulus interval above 3 seconds and also reported that it didn't sound like music (Gruber & Block 2013). This all fits well with earlier research arguing that 3secs is a time constant in perceptual tasks representing a central neural mechanism that functions to integrate successive events into a "gestalt" in order to create a "subjective present" (Pöppel, 1973, 1978). Much empirical evidence has, also, been collected to support the theoretical framework of the 'psychological present' in perception across modalities and in both perceptual and motor behaviour (see Wittmann & Pöppel, 1999/2000). It is no coincidence, of course, that music and melody provide constantly an ideal example in philosophical discourse on the binding of temporal experience (e.g. Husserl, 1977). It is within this temporal range that repetitive isochronous events have been shown to easily build perceptual expectancies towards future events (Jones, 1989; Large & Jones, 1999) and in which musical pulse is primarily operational (London, 2005; 2012).

Overlapping considerably with the previous time span is the notion of the "specious present" (James, 1890) or "living present" [lebendige Gegenwart] (Husserl, 1991). In the simplest sense this is a subjective experience of "now" of the "present moment". James' original estimate for this phenomenon was that it varied in length from a few seconds to "probably not more than minute" (p. 642). Block and Gruber's review concludes that subsequent research suggests a time interval of about **3 secs to, arguably, about 7 secs**, "during which the brain can compare and analyze very recent high density memories in working memory" (p. 130)[38]. James famously describes that "the practically cognized present is no knife-edge, but a saddle-back, with a certain breadth of its own on which we sit perched, and from which we look into two directions into time" (p. 609). Further on he states that "its content is in a constant flux, events dawning into its forward end as fast as they fade out of its rearward one … meanwhile the specious present, the intuited duration, stands permanent, like the rainbow on the waterfall, with its own quality unchanged by the events that stream through it" (p. 630). This idea that this "present" enables our awareness of change and sequence but is also itself a constant, invariable structural form of consciousness is also reflected in Husserl's characterisation of the standing present [nunc stans]. Husserl argues, however, that this is "flowing" rather than static. He concludes that "what abides, above all, is the formal structure of the flow. That is to say, the flowing is not only flowing throughout, but each phase has one and the same form" (1991, p. 118). As Dainton (2014, p. 391) summarises, his characterisation "combines sameness – an "absolutely abiding form" – with continual change and renewal… it is an invariant *structural* feature of our consciousness".

The main point with regard to the concerns of this thesis is that the range of time that generally equates to a single musical pulse is broadly experienced as a perceptual present. Therefore, by operating within a pulsed framework in music, we are constantly and continually joining one "moment" to the next (and with reference to the prior moment). The particular nature of period correction mechanisms in music will be discussed below to constitute an important psychological and phenomenological distinction between musical pulse and the general flow of experienced time.

Timing Mechanisms

It seems uncontroversial to assume that fundamental abilities to perceive and interact with the temporal structures of the physical world and to direct our movements in time are present in higher-order senescent animals. Leading researchers on comparative timing confidently state, for example, that "humans share with other animals an ability to measure the passage of physical time and subjectively experience a sense of time passing" (Allwood et al. 2014, p. 743) and that "an essential component of primate cognitive function is the ability to extract and represent temporal information from the environment. The quantification of the passage of time, in turn, is crucial to coordinate motor behavior" (Zarco, 2009, p. 3192). The

---

[37] His description of the types of temporal order and the time spans involved has provoked some discussion and refinement (see Block and Gruber, 2014). However, the shortest two have been accepted as being pretty clear and uncontroversial (see Eisler et al., 2008):

- 0-2ms: Within this temporal interval range simultaneous and non-simultaneous events are perceived to be simultaneous (Hosokawa et al., 1981; Moore, 1993).
- 2/3ms-20/40ms: Events separated by this temporal range create the impression of non-simultaneity but the temporal order of events - which of two events came first - cannot be confidently or unerringly distinguished (Hirsh, 1959; Lotze et al., 1999; Mills & Rollman, 1980; Steinbüchel et al., 1996; Pastore & Farrington, 1996).

It is important to note that these time-ranges seem to be operational across modalities, and in both perception and performance. Studies have shown, for example, that the threshold between events whose temporal order are or are not distinguishable is consistent for acoustic, visual and tactile stimuli (Hirsh & Sherrick, 1961). Furthermore, timing processes inherent in kinesthetic abilities is constant across a range of effectors (Zelaznik & Smith, 1992).

[38] Although the question of whether the processing of larger timing intervals is operated by similar or overlapping mechanisms is open (Gibbon et al., 1997), it is widely accepted that durations exceeding 3-7 seconds involve an additional memory process that links past moments with the present (Fraisse, 1984). This further timing mechanism is thought to bind successive events in perceptual units of 2 to 3 seconds (Block, 1990; Fraisse, 1984; Pöppel, 1978; 1997). Pöppel terms this process the formation of perceptual Gestalts. The wide-ranging literature on estimations of prospective and retrospective time intervals goes beyond the scope of this thesis (see Pöppel 1999).

nature of the central timing machinery – the ‘internal clock’ - has been the subject of a huge body of psychological modelling, experimental research, and debate (see Allman et al, 2014; Ivry & Schlerf, 2008; Grondin, 2014)[39]. It is particularly important to note that, although the literature has by and large supported a pacemaker-accumulator model, much of the focus of experimental tasks has been on the timing of single, isolated durations in human and non-human animals (Hills, 2003; Lejeune & Wearden, 2006).

Dynamic Attending Theory

An important model of timing, that has generally been viewed to have the most targeted relevance to relative timing and extended rhythmic contexts, is Jones’ Dynamic Attending model (Jones, 1976; Jones & Boltz, 1989; Large, 1994), in which internal oscillatory mechanisms create points of attentional focus and expectation (Large & Jones, 1999; Jones & McAuley, 2005). This theory is not so much focused on judgments of time as on the operational side of how we and other animals interact in time with our environments. Central to the theory is the view that attention is not a continuous operation but rather one in which we constantly build expectancies and subsequently direct efficient energy pulses towards expected events whilst reacting to the unexpected. A key feature of this model therefore is that it is attentionally *future-directed* with internal predictions based upon perceived regularities and/or learned patterns of events entraining with the external world. Large and Jones (1999, p.123) explain succinctly that:

> The goal-oriented mechanism reflects expectancy and attentional set; it is concerned with deliberate changes in focal attending (e.g., Folk, Remington & Johnston, 1992). For instance, an individual may react faster to an object in an attended-to-location or may be successfully instructed to ignore certain objects in a forthcoming display. By contrast, the stimulus-driven mechanism reflects fast (some claim automatic) attentional shifts, usually caused by a single salient feature or an abrupt display change. For instance, the abrupt onset of a non-target object may provoke a rapid attentional shift to it and away from a target object – (i.e. attentional capture) (Jonides & Yantis, 1988; Yantis & Hillstrom, 1994).

Differing models of Dynamic Attending Theory have been proposed (e.g. Large, 1994; McAuley 1995). A full discussion of the different approaches is beyond our current scope. However, they all share four fundamental concepts (see Jones 2016): 1) Neural oscillations are self-sustaining; generating inferred expectancies over time; 2) An oscillator has an intrinsically stable period that is only briefly disrupted by a violation of expectation; 3) Entraining oscillations can adaptively correct both its phase and period in response to violations of expectancy in a stimulus; 4) Multiple related oscillations are activated by multiple time levels within complex metrically and rhythmically structured events.

In recent years the dynamic attending theory framework has been shown to be an excellent theoretical fit with results of research into neural entrainment. Stimulus events occurring at excitable phases of entrained neural oscillations are, for example, more likely to be detected or be responded to more quickly (Cravo et al., 2013; Schroeder and Lakatos, 2009; Stefanics et al., 2010). Henry and Herrmann (2014) suggest specifically that oscillations in the excitable phase of low-frequency neural oscillations in the delta-theta range might reflect correlates of the attentional oscillation with peaks correlating with the attentional pulse in models of dynamic attending theory. Interestingly from a comparative perspective “rhythmic” sensory stimuli have been shown to correlate with entrainment of brain oscillations in humans, macaque monkeys, and zebrafish (Lakatos et al, 2008; Saleh, et al, 2010; Sumbre, et al, 2008). Wilson and Cook (2016, p. 3) note that “in macaques and humans, at least, this propagation of timing extends up to and includes the motor system”. This is evidenced by decreased reaction times under conditions of rhythmic input (Lakatos et al., 2008; Praamstra et al, 2006; Saleh et al., 2010).

Dynamic Attending Theory has been associated particularly with music as it naturally resonates with models of musical time (see Jones 2016) and offers a clear explanatory model for the psychology of pulse in music. The generically increased levels of predictability and stability in music have also made it an ideal experimental testing ground and empirical paradigm for the theory. Nevertheless, it is worth noting that the theory in itself seeks to explain our broad capacities to interact with the temporalities of our movements and environments rather than music specifically. A key point, therefore, is that a musical pulse is not intrinsically separate but rather is built upon more generalisable mechanisms that are employed in a specific way (in particular being framed by attentional correction mechanisms discussed below). These different models naturally suggest some key questions for the concerns of a comparative perspective on music: Are beat perception and interval timing facilitated by different neural circuits? And are they constituent processes of a wider internal clock, or are they two divergent timing systems?

---

[39] Incorporating arguments for shared mechanisms across perception and performance, the search for an internal clock has largely been dominated by the sensory-accumulator model in which pulses that are emitted regularly by a pacemaker are temporarily stored in an accumulator (Woodrow, 1930; Treisman, 1963; Gibbon et al., 1984; Ivry and Richardson, 1995; Meck, 2003; see Buhusi & Meck, 2005). Over the course of the last decade in particular the emphasis has shifted towards understanding the neural basis of and mechanism for interval timing (see Gibbon et al., 1997; Harrington et al., 2004; Ivry and Spencer, 2004; Merchant et al. 2013; Bahusi & Meck 2005).

Beat-Based Versus Duration-Based Timing

The following section discusses evidence showing a clear neurophysiological distinction between beat-based and duration-based timing in humans. I will also show that relative (beat-based) timing is a clear and critical difference between our species and other primate species who appear to operate almost exclusively through an understanding of absolute timings. Finally, I present three evolutionary hypotheses (Patel & Iversen, 2014; Large et al., 2015; Honing & Merchant, 2013) for the emergence of beat-based timing in humans arguing that they are complementary rather than being mutually exclusive.

By and large perceiving and producing musical rhythms are strongly associated with areas of the brain involved in motor function (see Cameron and Grahn, 2015; see below). Neuroscientific and neuropsychological studies seem also to strongly suggest a difference between beat-based and non-beat based timing. In response to rhythms that induce a beat compared to those that do not, the basal ganglia and supplementary motor areas are more active (Grahn & Brett, 2007; Grahn & Rowe, 2009; Teki et al., 2011). Conversely, during absolute timing perception and active tasks the cerebellum is more active (Grube et al. 2010; Teki et al. 2011). Interestingly, in a continuous tapping task comparing beat-based and non-beat based rhythms, the basal ganglia's response profile suggests a role in beat prediction, but not in beat finding (Grahn & Rowe, 2013). However, Kung et al. (2013) found the opposite with a different task requirement in which the contrast was between the initial phase of active tapping to a beat-based rhythm and listening to an isochronous pattern. Thus it seems plausible that the basal ganglia are involved in both beat finding and continuation depending on the task requirements and/or experimental controls used.

Dissociations are also evidenced in neuropsychological work on patients with Parkinson's disease. Notably this disease is partially indicated by compromised basal ganglia function. Patients consistently show selective deficit in discriminating changes in beat-based rhythms but not non-beat based rhythm (Grahn and Brett, 2009). Rhythmic auditory stimulation (e.g. Thaut et al. 2015) has also been shown to be a highly successful therapeutic tool in Parkinson's patients and will be discussed further below. A consistent contrast is that patients with cerebellar degeneration show a deficit for absolute rather than relative timing (Grube et al., 2010a).

Further support for a functional distinction comes from FMRI studies of processing material that can be interpreted either relatively or in absolute terms. McAuley et al. (2006) devised an interestingly ambiguous paradigmatic stimulus in which a periodic (600ms) beat is implied but not explicitly emphasized. This leads to contrasting reports from participants of either slowing down or speeding up at the end of a row of tones, thus indicating a difference in whether the sequence was processed in relation to the implied beat or in terms of more absolute interstimulus intervals. Crucially both processes were equally effective as researchers found no evidence that duration-based timing was less accurate or precise than beat-based timing. Subsequent neuroscientific analysis (Grahn & McAuley 2009) supports the idea that different strategies are supported by cortical activation differences reflecting the engagement of different neural timing mechanisms. The authors conclude (p. 1901) "that individual differences in behavioral bias toward beat perception correspond to activation differences in left cortical motor areas and insula, whereas the opposite bias is correlated with right premotor and posterior auditory cortex activation."

Overall it seems that, in the range of timings relevant to beat-based processing/entrainment, neuroscientific research supports the idea that the cerebellum is connected with absolute timing of subsecond intervals (Diedrichsen et al. 2003; Grube et al. 2010a,b; Teki et al. 2011), and that corticostriatal circuits are more operational in the relative timing of suprasecond intervals and in beat-based perception and production (Coull et al. 2011; Grahn & Brett 2009; Grahn & McAuley 2009; Grahn & Rowe 2012; Meck 2005; Schwartze et al. 2011, 2012a,b; Teki et al. 2011).

Absolute vs relative timing in primates

It seems likely that duration-based timing mechanisms are widely shared amongst primate species but that beat-based timing is either unique or particularly advanced by many degrees in humans. An important study addressing this issue was carried out by Zarco et al. (2009). They showed that (after extensive training [11-25months!]) three Rhesus Macaques were able to produce coordinated tapping movements, matching a range of randomly presented tempi (450-1000ms), during a synchronization–continuation task (SCT). Analysis was collected from brief bouts in which the animals produced six isochronous intervals in a sequence, three guided by stimuli, followed by three internally timed (without the sound) intervals The macaques generally reproduced the intervals with slight underestimations in the region of 50ms and the variability between inter-tap intervals generally increased, as it does in humans (see below) in positive correlation to the target interval. Donnet et al (2014) have subsequently shown that these monkeys use an explicit timing strategy to perform the SCT, similar to humans (Doumas & Wing 2007), in which they produce isochronous rhythmic movements by temporalizing the pause between movements and not the movements' duration. Unlike human sensorimotor synchronization (see below), however, the taps of macaques typically occur about 250 ms after stimulus onset. Thus while some predictive mechanism is suggested in that the figure is lower than expected reflex reactions[40], it strongly suggests that the monkeys were not using an analogous beat-based strategy to humans. In the strict sense of the term they were not entrained and were relying primarily on absolute rather than relative cues. EEG studies on rhesus monkeys also tentatively support this distinction showing MMN (mismatch negativities) for violations of simple regularities but not to a musical beat (Honing et al, 2012). Another interesting

---

[40] Merchant and Honing (2014) describe some unpublished effort to lower this positive mean asynchrony following further incentives and claim to have been able to reduce to a following response to roughly 100ms.

difference to human participants on similar trials (see below) is that the macaques showed no bias toward auditory as opposed to visual cues for synchronization.

Some other studies have reported sensorimotor synchronization in chimpanzees and bonobos. However, while the studies are relevant and perhaps indicative of some form of oscillatory entraining mechanism, the details are less impressive than the article titles might suggest. Hattori et al (2013) describe spontaneous entrainment to an auditory rhythm in a chimpanzee using keyboard keys. However, only one of three animals showed any signs of entrainment and exclusively in short bursts when the metronomic so-called “distractor sound” was at 600ms, near the animals preferred rate of 579ms. In other words it showed brief phase alignments but no evidence of consistent period correction (see below). Large and Gray (2015) also report ‘spontaneous tempo and rhythmic entrainment in a Bonobo (*Pan paniscus*)’ – ‘Kuni’. Whilst they analyze and report that entrainment episodes were highly unlikely to have occurred by chance, only 13% of 214 episodes of simultaneous drumming in which Kuni produced at least 12 contiguous drum strikes were significantly entrained with the metronome. Although, the authors used different tempi they all were very close to Kuni’s preferred (and fast!) tempo of 270bpm – (230bpm – 280bpm). Once again it seems that the bonobo exhibited a degree of occasional oscillatory phase alignment rather than any form of consistent attentional entrainment and error correction[41].

The current evidence does seem to suggest some impressive absolute timing capabilities in primates similar to our own, but, by and large, a lack of beat-based or relative processing. Darwin most likely was at best partially right when he wrote that “the perception, if not the enjoyment, of musical cadences and of rhythm is probably common to all animals, and no doubt depends on the common physiological nature of their nervous systems” (Darwin, 1871, p. 174).

Three hypotheses on the evolution of beat-based timing in humans

Patel & Iverson’s (2014) suggestions for the action simulation for auditory prediction (ASAP) hypothesis describe that the motor planning system uses “a simulation of body movement (specifically, of periodic movement patterns) to entrain its neural activity patterns to the beat period, and that these patterns are communicated from motor planning regions to auditory regions where they serve as a predictive signal for the timing of upcoming beats and shape the perceptual interpretation of rhythms” (p. 1). They also make specific neuroanatomical predictions that the neural substrates of beat perception involve the dorsal auditory pathway, and that this might be a possible locus of neuroanatomical differences between humans and monkeys. Neuroanatomically this may prove to be an interesting evolutionary suggestion, especially as it may also relate to vocal learning in our recent evolutionary past. However, the ASAP theory, as it stands, suffers from the fact that the music or beat appears, here, to exist physically before it is perceived. Music is seemingly being described as an autonomous physical product which, particularly in an evolutionary context, is problematic. In this respect the theory of neural resonance may provide an alternative understanding of the nature of the processes which operate both in terms of perception and performance.

In contrast to the ASAP hypothesis, the neural resonance approach (Large et al., 2015) makes specific predictions about the nature of the neurodynamic processes involved in musical rhythm perception (see Large, 2008; Large and Snyder, 2009) rather than the neuroanatomical locus. It claims, in essence, a more gradual link between human and primate beat-based timing abilities arguing that “an oscillatory network interaction can lead to spontaneous pulse induction in complex rhythms—even in the most extreme case of a rhythm for which there is no energy at the pulse frequency (Large, 2010; Velasco and Large, 2011)” (p. 7) This evolutionary possibility would seem to correlate with local field potential (LFP) studies in macaque monkeys showing beta- and gamma- band oscillations measured during synchronization-continuation tasks (Bartolo et al., 2014) and observations of beta oscillations in the basal ganglia during synchronization and continuation (Bartolo and Merchant, 2015; Merchant & Bartolo, 2017; see Merchant et al., 2015a). Large and colleagues explain that these suggest differential roles in rhythmic tapping versus stimulus processing and further support their role in the striato-thalamocortical circuit during control of rhythm. Activation of cell populations in the premotor cortex of rhesus macaques that appear to be stimulus-predicting cells also appear to respond to regularly timed visual or auditory stimuli (Merchant et al., 2015). The neural resonance approach currently falls short, however, of providing a definite interpretation of the differences between human and primate timing capacities. They themselves acknowledge this and tentatively suggest that “species differences could be accommodated within a neural resonance framework, for example, simply in terms of the differences in auditory-motor coupling (see e.g., Merchant and Honing, 2013)” (p. 10).

Merchant & Honing’s gradual audiomotor evolution hypothesis suggests that rhythmic entrainment (or beat-based timing) developed gradually in primates, peaking in humans but is present only with limited properties in other non-human primates. The approach highlights that macaques show equivalent temporal performance to humans in tasks that involve the perception and production of single intervals. For example, the authors reference Mendez et al., (2011) who showed that the relative psychometric threshold for the categorization of single intervals given by brief visual cues is very similar between humans and macaques. They argue that as, in humans, audition dominates temporal processing (Repp and Penel, 2002; Bertelson and Aschersleben, 2003), it is quite possible that the human auditory system has advantaged contact to the temporal and sequential mechanisms working inside the mCBGT circuit and that this partially determines the excellent rhythmic abilities of our species. This, as the authors, explain matches suggestions by Rauschecker and Scott (2009) that

---

[41] In a recent interactional, but not necessarily “pulsed”, tapping study with chimpanzees one of four chimpanzees (Chloe) did produce significant convergence in both tapping tempo and timing towards those of a conspecific partner in auditory and face-to-face settings (Yu & Tomonaga, 2015; 2016). Even with this individual, however, synchronization was relatively weak compared to that of humans and it is not entirely clear to what degree her performance was due to specifically relative processing.

"the privileged access of the humans' auditory system to the sequential and temporal machinery of the mCBGT circuit emerged gradually in the course of evolution from precursors of the great ape lineage" (p.5). They also cite further support from a study showing that the connectivity between the superior temporal auditory areas and the frontal lobe in chimpanzees has an intermediate level of complexity when compared with macaques and humans (Rilling et al., 2008; Petrides and Pandya, 2009). It would be premature to try to conclude anything definite at this stage. However, all three hypotheses are to be commended on offering clear predictions and potential routes for future experimental support or falsification. This is clearly an area that will be explored further in the next few years.

Other less closely related species have been argued to exhibit entrainment capabilities (see below) but, despite some understandable excitement and interest, these need to be understood comparatively in terms of whether or not they represent a true analogy to how sensorimotor synchronization is psychologically manifest in our species. Sensorimotor synchronization has been extensively studied for over a century (Stevens, 1886; see Repp 2005; 2013) predominantly in an experimental paradigm in which participants tap to match a range of metronomic and/or perturbed stimuli. The research and experimental findings are of course indicative of *some* of the mechanisms involved in temporal coordination in music perception and performance. We should be clear, however, that a full understanding of interpersonal musical entrainment (Bispham, 2003; Clayton et al 2005) in interactive and social contexts requires an additional understanding of our capacities for culture action-mirroring and intersubjectivity (see Himberg 2014; Keller et al. 2015; Nowicki et al. 2013; see below).

Sensorimotor Synchronisation in Humans

Sensorimotor synchronization describes behaviours in which we match our actions to an external source based upon generated expectations as to the timing of future events. In its simplest form it describes a simple task to tap with a metronomic beat – a task we universally find to be straightforward and almost irresistible. Before moving on to tasks in which participants match an external source, the next section describes experimental analyses of participants producing a simple periodic pulse. Attentional mechanisms that afford this capacity provide a clear psychological, and behaviourally observable, difference between this and instances in which the biomechanical nature of a particular action in humans and/or other animals produces a reasonably steady "beat". In the following section I will present hypotheses as to how we perceive synchronized action and why, in sensorimotor synchronization tasks, we tend to tap asynchronously slightly ahead of the beat. Understanding this is crucial to a comparative analysis of whether capacities exhibited in other species are truly analogous to sensorimotor synchronization in humans. Also crucial in this respect is a full understanding of the correction mechanisms involved in maintaining coordination with an external source and/or interactive partner. I will, therefore, also describe phase and period correction mechanisms, the latter being more cognitively challenging, subject to a degree of awareness and volitional control, and arguably specific, amongst communicative settings, to a musical pulse. Finally, I will present analysis of documented cases of sensorimotor synchronization in the animal kingdom arguing, with reference to the cognitive theories and models in the preceding sections, that none provide a full psychological and behavioural analogy to sensorimotor synchronization in humans.

Periodical production

The pioneering study into isochronous serial interval production (ISIP) was run by Stevens in 1886. Participants were instructed to continue a metronomically established pulse (360ms-1500ms) and the analysis established some important findings. Subjects, predictably, had no problems adjusting to the response rate of the metronome beat and could maintain it over extended periods. However, this study showed considerable variability in performance and an increased variability with increases in temporal interval. Stevens (1886) observed from a visual inspection of the data that variability between the target interval and the mean of the intervals produced by the subjects remained within about 5% of the standard deviation and that variability was evident on two distinct levels (Ibid). Short-term fluctuations were characterised by a zig-zag shape such that two sequent variations in the same direction were rare and longer waves extending over the sequence (subsequently termed 'drift').

Inter-onset-interval variability

Subsequent studies have confirmed a typical standard deviation of ISIP series of between 3-6 percent of the inter-onset-interval (IOI)[42] with a variety of experimental controls and tasks affecting the levels of dispersion to some degree (e.g. Keele et al., 1985; Madison, 1998; Michon, 1967; Wing & Kristofferson, 1973b). A training effect has been evidenced with feedback reducing dispersion by thirty percent (Nagasaki, 1990). Nevertheless, dispersion remains reasonably substantial even in the case of extensive rhythmic training (Madison, 1992) and indications as to whether musicians produce less dispersion is mixed with some studies supporting improved periodicity (Franêk et al., 1991; Keele et al,. 1985) and others showing no difference (Madison, 1999; Yamada & Tsumura, 1998). The increased interval to interval deviations with increases in IOI observed by Stevens as well as the 'zig-zag shape' – whereby an interval longer than the mean is frequently followed by an interval shorter than the mean – have also been further validated (Vorberg & Hambuch, 1984; Vorberg & Wing, 1996; Wing & Kristofferson, 1973a; 1973b). The former has been explained as an accelerating function of the IOI (Fraisse, 1984) with breaks around 250ms (Wing & Kristofferson, 1973b) and 1 sec (Madison, 1998). Madison (2000, p. 86) has argued that 'the latter break is probably related to qualitative changes in performance found for other temporal tasks' (Madison, 1998; Szelag, 1997). These, he posits, may be related to a cross-modal retention or integration limit around one

---

[42] Inter-onset-interval simply refers to the time between taps. It is effectively a measure of the speed/tempo of the metronome or beat.

second, in auditory (Crowder, 1993; Kubovy & Howard, 1976; Mates et al., 1994; Pöppel, 1996; 1997; Warren, 1993) and visual processing (Freyd & Johnson, 1987).

The 'zig-zag' negative first-order correlation observation has been influentially modelled mathematically by Wing and Kristofferson (1973). They supposed that a series of timed pulses are generated by an internal timekeeper. In their model each internally generated interval is subject to a delay in motor implementation before the occurrence of an observable response. Internally generated intervals and motor implementation delay are seen to be independent and subject to random variability such that, over successive responses, the observed lag-one autocorrelation is predicted. Importantly, the model suggests that this occurs as a result of the two-level architecture with distinct and behaviourally dissociable random fluctuations in both timekeeper and motor implementation and not from some form of feedback mechanism. Numerous studies each emphasising different variables such as the age of the participants, levels of 'clumsiness', and time perception as well as using a variety of effectors have supported the model (see Wing, 2002). They have unsurprisingly shown that increased variance with increases in inter-tap-intervals result from variances in the internal timekeeper (Wing 1980). Typically, relatively short sequences have been used in support of the model in order to avoid the factor of 'drift' and efforts have been advanced to adapt the model to make explicit allowance for resultant estimator bias (e.g. Vorberg & Wing, 1996). Drift refers to involuntary substantial fluctuations in the mean proportional to the IOI over time – essentially a form of subconscious accelerando and ritardando. It constitutes a positive higher-order dependency as several intervals in succession become increasingly shortened or lengthened. Madison (1998; 2001) has notably attempted to incorporate the extent to which drift contributes to variability in isochronous tapping tasks. His model of variability accounts for and distinguishes between negative first-order correlation and drift by subtracting a linear marked trend in a sequence from total dispersion and hence calculating the percentage of mean dispersion caused by drift. Without going deeply into the mathematics employed (see Madison 2001 for details) it is clear that, even after subtraction of the trend (drift direction), the remaining variability retains positive higher-order dependence as shown by a 5 point moving average.

We have effectively evolved an operationally cross-domain timing system that is *sufficient* for building expectancy and coordinating internal and external events, but without a need for absolute precision. Crucially, the implication of an internal timekeeper model is that this experimental paradigm engages a process that deals with psychologically manifest timing mechanisms (partially shared with other primate species, see above) *between* actions (taps) rather than periodicities occurring naturally as a result of particular movement patterns[43]. Wallin et al. (2000) introducing the influential collection of chapters in *The Origins of Music* state that "most animals (including humans) have the ability to move in a metric, alternating fashion" (p. 11). They contend that "what is special about humans is not their capacity to move rhythmically but their ability to entrain their movements to an external timekeeper, such as beating a drum" (p. 12). This proposal will be debated further below but the studies of continuation tapping ('periodic production') show that, even lacking an external source to match, there are psychological models that we can subject to comparative analysis. The most important point to take from the above discussion is that periodic production in nonhuman species can only be considered as being homologous (and of direct relevance) to these periodic processes if its behavioural correlates are not explicable solely in terms of biomechanical efficiencies. A "musical" isochronous or "rhythmic" production implies reference to an external timekeeper *and/or* to an internally created and, in part, volitionally controlled attentional pulse (e.g., Drake & Bertrand, 2001).

"Drumming" – in which animals strike repeatedly on hollow objects within their natural environment in communicative, territorial or courtship contexts - has variously been reported in many species including palm cockatoos (Wood, 1988), woodpeckers (Dodenhoff, Stark, & Johnson, 2001), and kangaroo rats (Randall, 1997). In primates we have examples of bimanual drumming in gorillas (Schaller, 1963), chimpanzees (Goodall, 1986), bonobos (unpublished data mentioned in Fitch, 2006), and macaques (Remedios et al. 2009). These are potentially of great relevance due to the relative phylogenetic proximity and possible homologies to human drumming in terms of the motor skills employed. In chimpanzees, drumming on buttress roots is exhibited predominantly by the dominant male in conjunction with displays of vocal pant-hooting. Communicative intent is suggested by aspects of the drumming being correlated with resting periods and the future direction of travel in group foraging. The displays have also been shown to be both individually and culturally distinct (Arcadi et al., 1998; 2004). Remedios et al. (2009), observing social macaque groups, found that these animals also use artificial objects to produce salient "periodic" sounds. They argue, following behavioural tests and a preferential looking analysis, that these drumming sounds attract the attention of listening monkeys similarly to conspecific vocalizations and that drumming sounds influenced the way monkeys viewed their conspecifics. Their suggestion is that drumming serves as a multimodal signal of social dominance. In an interesting neurological follow-on experiment (Remedios, 2010), they report further that the representations of both drumming and vocal sounds overlap in the caudal auditory cortex and the amygdala thus supporting a common origin of primate vocal and nonvocal communication systems and possibly by extension a gestural origin of speech and music. However, in contrast to the human tapping experiments described above there is, to my knowledge, no evidence to suggest that these "drumming" episodes are motivated towards isochrony or that any resultant regularity goes beyond those inherent in the biomechanics of the actions involved.

Another reported instance of the production of a steady beat in the animal kingdom is Asian elephants in northern Thailand and their participation in the 'Thai Elephant Orchestra'. These elephants have been trained to strike percussion instruments. Baroni et al (2006) quantified the mean and standard deviation of the inter-beat-interval for eleven drumming sequence and found that the elephant's drumming tempo was remarkably stable across sequences, ranging from 33.4 to 36.1 beats per minute (mean = 34.9, sd = 0.9). Temporal variability, quantified as the coefficient of variation of IBI within each sequence, averaged 3.5%, which, as the authors note, is lower than the variability found when humans tap at this tempo (typically 7%).

However, there was no evidence of the elephants adjusting to different tempi or efforts to match external stimuli. It seems plausible, therefore, that this also reflects merely the animals' preferred rate for a particular biomechanical action rather than an output with reference to a sustained attentionally driven pulse

Sensorimotor synchronisation - A universal human skill

It is crucial to note, in stark contrast to the studies of animal "entrainment" (see below), that this appears to be an absolutely universal skill in humans. Individuals categorized as being "amusic" (see Peretz et al, 2002; Ayotte et al 2003; Peretz 2006) show fine-grained pitch perception difficulties (see chapter 3) but show a normal ability to synchronize movement to the beat of popular dance music as well as potential for improvement when given a modest amount of practice (Philips-Silver et al, 2013). In the one reported case of "beat-deafness" the individual, in fact, showed near-normal synchronization with a metronome suggesting that his deficit was concerned with finding the beat in music (Philips-Silver et al, 2011) rather than sensorimotor synchronization per se. Typically this ability is developed by the age of 4-6 with age-specific entrainment regions evidently narrower in childhood and late adulthood than in midlife (see McAuley et al, 2006). ERP studies would appear to further suggest that some form of beat induction and corresponding future expectation is innately neurologically manifest (Winkler et al. 2009).

In the much explored experimental paradigm of tapping to an isochronous and/or perturbed metronomic stimulus, participants naturally exhibit considerably reduced variability (Dunlap, 1910). In some ways, presumably reflecting some of the difficulties in self-pacing, this task appears to be neurologically *less* complex than free continuation tapping (Rao et al, 1997; Gerloff et al, 1998 Miall, 2004 cf. Jantzen et al 2004). However, it is also necessarily dependent upon additional error correction mechanisms to counteract individual timing and motor variabilities and maintain perceived synchronisation (Mates, 1994; Schulze & Vorberg, 2001; Semjen et al., 1998). Assuming that the inter-onset-interval (IOI) is set between 100-170msec and roughly 1.8 seconds participants generally have no problems with this task. It is important to note that the lower IOI limit seems to relate to an auditory integration window estimated in studies of the mismatch negativity (e.g. Yabe, et al 1997; Näätänen 1998; Yabe et al., 1998). Musically trained individuals appear to be able to synchronise at lower IOI settings and may correspondingly, as Repp (2005) suggests, have particularly narrow integration windows allowing optimal temporal resolution of events. It is also interesting to note that the lower synchronization threshold is about four times as high for stationary visual stimuli as it is for auditory stimuli (Repp, 2003). This lower IOI limit also seems to correspond in music to the shortest duration of tones that can operate as individual elements in a musical rhythm (Friberg & Sundström, 2002; London, 2002, 2012). Generally speaking tones that are shorter than 100ms tend to be perceived as being ornamental in function. Repp (2005 p. 973) summarises that the lower and upper IOI limits are musically relevant in that they coincide approximately with "those within which a sequence of events can be perceived to have rhythmic and metrical structure (Bolton, 1894; Fraisse, 1982; London, 2004)". Beyond these limits additional mechanisms or techniques of psychological amalgamation or subdivision are putatively assumed to be operational (see London 2012 for further discussion). Crucially in terms of some of the later conclusions in this thesis the upper IOI limit can be seen to be connected to the perceptual present and the temporal capacity of working memory (see above, Pöppel, 1997; Szelag et al., 1996; Wittmann & Pöppel, 1999/2000).

The perception of sensorimotor synchronisation – evidence from asynchronies

One of the most critical and consistent features of sensorimotor synchronisation is that we tend to tap prior to the stimulus onset. This is commonly termed 'negative mean asynchrony' (NMA) and presents evidence of a predictive mechanism and stable entrainment. Although our taps precede the stimulus, we will experience our actions as being precisely aligned. The reasons for this perception of synchrony has been the subject of much debate and varied hypotheses[44].

---

[44] The earliest suggestion was the **Nerve-Conduction (Paillard-Fraisse) Hypothesis,** in which the origins of NMA are seen to be in the integration peripheral processes. The tactile and kinaesthetic feedback to the brain from the tap is argued to be slower than the auditory sensory information. Hence the theory is that the tap has to precede the click to establish synchrony at the level of central representation (Aschersleben & Prinz, 1995; 1997; Fraisse, 1980; Paillard, 1949). Supporting this hypothesis is evidence of increased NMA in foot-tapping versus finger-tapping (Aschersleben & Prinz, 1995; Billon et al., 1996) and in studies that artificially increase and/or delay auditory feedback (Ascherleben, 1995, 1997; Ascherleben & Prinz, 2007; Mates & Aschersleben, 2000; O'Boyle & Clarke, 1996; Mates et al., 1992). Considering the varying quantitative synaptic transmissions involved, this model can also accommodate effects of the pacing signal's sensory modality on the magnitude of NMA (Kolers and Brewster, 1985; Müller et al. 2001). Significantly, however, the hypothesis cannot accommodate results showing the order of NMA produced by different effecting organs nor results indicating effects of practice (Aschersleben, 1994; 2000a; 2000b); musical expertise (Ludwig, 1992; Repp and Doggett; Repp, 1999; Krause et al 2010; Fujii et al. 2011); changes in the duration of the tap (Vos et al 1995); changes in the force of the tap (Aschersleben et al., 2003a; Gehrke, 1996); or the nature of the inter-interval space (Thaut 1997; Repp, London & Keller 2005; 2008; 2011; 2013).

The **Perceptual (P)-centre hypothesis** proposes that perceived synchronisation occurs not at the onsets of tap and pacing signal but rather at their respective perceptual centres. Effectively the longer the duration of the stimulus, the more the p-centre is shifted away from the stimulus onset. Thus the hypothesis accords with a delayed p-centre in finger tapping versus the rather curt auditory clicks that are commonly used (Vos et al., 1995). This hypothesis, however, fails to account for all the effects on NMA outside of the duration of the pacing signal.

The **Sensory Accumulator Model** supposes that an external event has to be experienced as one neural entity. As such it shares with the nerve-conduction hypothesis the notion that synchrony is fundamentally established at the level of central representation. In contrast, however, it is the density of afferent neural signals generated by the physical events that determine the processing times necessary for the creation of the neural state. The model describes this as an 'accumulation function' that determines the time elapsed between an external

It seems fair to say, and the view of the leading figures in this research area, that the origin of the NMA in these tasks remains unclear with none of the models in full agreement with all the available evidence. Nevertheless, this body of investigation together with the analyses of putative correction mechanisms (see below) does capture *some* of the mechanisms implicit in more naturalistic human pulse-based sensorimotor coordination and thus offers some grounds from comparative scrutiny in other animals (see below). A key point to make is that none of the available models suggest anything that is necessarily unique to a musical pulse. Rather, they describe more general mechanisms and processes through which we perceive events and our own actions as being synchronized. Addtitionally, the theoretical models and the behavioural evidence of NMA is crucial in comparing putatively analogous behaviours in other species. We should, however, also be very clear that they by no means represent the whole picture of how sustained periodically structured coordination is obtained and sustained in social and/or musical settings. It is necessarily the case that in an interactive musical setting the mean asynchrony in relation to an interactively generated and shared pulse between all members of a dyadic or ensemble performance must in effect approach zero. Part of this can be discussed in terms of psychological and kinesthetic training effects in musicians. For example, Stoklasa et al, (2012) reported that musicians playing their own brass or string instrument showed an almost neglible NMA when synchronising to a metronome (–2 ms) (tapping (–13 ms)). However, even in pared down tapping studies of 'social entrainment' we seem to rely also on a degree of interpersonal resonance, action empathy and/or other social cues. Himberg (2014), for example, showed that, compared to tapping with a 'deadpan' metronomic signal or a 'phasevar' computer program (mimicking 'human' variability), two humans displayed the highest degree of synchronisation despite higher levels of drift and "instability". These results, as Himberg (Ibid) states: "are strong indications that the "information-theoretic" concepts and experiments on time-keeping and synchronisation only cover a limited range of the spectrum of human timing and entrainment abilities" (p. 176). The social nature of a musical pulse will be discussed further below.

Error correction in isochronous and perturbed synchronisation

Sustaining regular sensorimotor synchronisation, even with an artificially isochronous pulse, requires some form of corrective mechanisms. Without these, timing errors due to motor and/or internal timekeeper variance (Wing & Kristofferson, 1973) would simply accumulate and lead to a loss of synchrony (Hary & Moore 1985; Vorberg & Wing, 1996). Of course in tapping with another human or in a musical setting we are subject to further voluntary and involuntary fluctuations due to features of individual style (Collier & Collier, 1996; 2002)) as well as expressive and structurally motivated modulations of tempo and microtimings (Schaffer, 1981, 1982; Gabrielsson, 1986; Palmer, 1989; Repp, 1992; Iyer, 2002). Studies using experimentally controlled perturbations to the target stimuli have provided a wealth of data and have been interpreted with regard to two principal models of error correction mechanisms: Dynamic systems theory and information-processing theory. In both cases they represent, respectively, extensions of more general models of timing described above. Dynamic systems theory assumes, in accordance with Jones' dynamic attending theory (see above) that an external rhythmic signal evokes intrinsic neural oscillations that entrain to periodicities in the rhythmic sequence (Large, 2008). As such, the mathematical modelling of corrections supposes non-linearly coupled oscillators that are described formally in terms of differential equations (e.g., Schöner and Kelso, 1988; Fink et al., 2000; Assisi et al., 2005). The information-processing theory assumes, instead, the principles of timekeeper models. Hence error correction is described according to linear autoregressive processes (e.g. Vorberg and Wing, 1996). In general, deviations of the current tap from the mean inter-tap interval and mean asynchrony is seen to be proportionally related to the deviation of the inter-tap interval and asynchrony associated with the previous tap plus random noise (Wing and Beek, 2002). It is possible that the two approaches account for different synchronization processes with each, therefore, more closely describing distinct aspects of the task (Torre and Balasubramaniam, 2009) or it may be perhaps more parsimonious to assume that both function as a general control equation for referential behavior (Pressing, 1999). Loehr et al (2011) argue that their results of sensorimotor synchronization analysis in piano playing preferentially fit a dynamic model. However, the particular choice of linear model (Schulze et al 2005) has been questioned given the complex nature of the music (see Repp and Su, 2013). In general, the majority of the experimental research projects have assumed an information-processing approach, although increasing evidence of oscillatory processes at the neural level (see Large et al 2015) perhaps suggests that a dynamic systems approach may ultimately represent a closer analogy to the true psychological processes involved.

---

event and its central representation. The steepness of the accumulation function and corresponding processing time in turn depends on the density of afferent signals. As such the model predicts that in a standard sensorimotor synchronisation task the NMA results from a faster accumulation of the auditory information than the tactile/kinesthetic information (Aschersleben et al, 2003a; Gehrke, 1995; 1996). This model accommodates studies of NMA with different body parts (Wohlschläger & Koch, 2000), modifications to the amplitude and force of the tapping motion (Aschersleben et al., 2003a; Gehrke, 1996), and effects of training and musical expertise (Aschersleben, 1994; 2000a; 2000b; Ludwig, 1992; Repp and Doggett, 2007; Repp, 1999; Krause et al 2010; Fujii et al. 2011). Bialunska et al. (2011) however, showed, contrary to the predictions of the model, that varying stimulus intensity affected simple reaction times to unpredictable stimuli but had no effect on the NMA.

Finally, Wohlschläger & Koch (2000) argue that NMA results from **errors in time perception**. They hypothesised that we structure time with the aid of auditory signals and bodily movements and that the simplicity of the production motions and emptiness of the IOIs employed in experimental settings lead to errors. This idea is again only partially supported by the diverse studies. It does fit with certain results following alterations to the kinesthetic nature of the tap (Aschersleben et al., 2003a; Gehrke, 1996). However, studies of subdividing the IOI with a stable inter tap interval have produced mixed evidence (with some suggestions of a musical training effect) (Flach, 2005; Zendel, et al, 2011; Repp 2008; Loehr and Palmer, 2011; see Repp and Su 2013). The predictions also fail to fully encapsulate the effects on NMA from changes to internal and external sensory feedback (Ascherleben, 1995; O'Boyle & Clarke, 1996; Aschersleben & Prinz, 1997; Mates & Aschersleben, 2000; Aschersleben et al., 2001; Cooke, et al, 1985; Cole & Sedgwick, 1992; Aschersleben, et al, 2003b), or from changes to the duration of the pacing signal (Vos et al., 1995).

Phase correction and period correction

In both models two interacting correction mechanisms are widely accepted to be independently operational: phase correction and period correction ((Mates, 1994a, b; Repp, 2000, 2001; Semjen et al., 1998; Vorberg & Wing, 1996; see Repp, 2005). Phase correction essentially adjusts for asynchronies between the last response and stimulus events assuming an unchanged period. Period correction instead modifies the next target interval on the basis of discrepancies between the oscillatory or timekeeper interval and the last or last few interstimulus intervals, thus altering the period of the attentional musical pulse. Phase and period correction are seen to be behaviourally and neurologically distinct. A combination of fMRI, EEG and TMS studies have documented wide subcortical and cortical networks incorporating the cerebellum, basal ganglia, premotor cortex, supplementary motor area, sensorimotor cortex, superior temporal gyrus, and inferior frontal gyrus (see Witt et al., 2008). Praamstra et al (2003) attempted to explicitly identify the distinction between phase and period correction in an EEG study that used source localization. They positioned the former in auditory and secondary somatosensory cortices and the latter in medial frontal cortex, in particular the supplementary motor area. Different patterns of functional connectivity have also been shown depending on whether corrections are automatic or effortful (Rao et al., 1997; Jäncke et al., 2000; Oullier et al., 2005; Chen et al., 2008; Thaut et al., 2009; Bijsterbosch et al., 2011a,b). This putatively corresponds to phase and period correction respectively (see below) and the functional distinction has noticeably also received support from an fMRI study of SMS with virtual partners (Fairhurst et al., 2012).

According to an information-processing approach phase correction mechanisms are thought to adjust for phase errors between stimulus and response events by taking the consciously or non-consciously perceived phase error (e.g. asynchrony) between the last response and stimulus events as input and subsequently correcting the next internal timekeeper interval by increasing or decreasing it by some fraction of that error (for a discussion of the nature of the mechanism see Schulze & Vorberg, 2001). A phase correction mechanism is generally believed to be sufficient for compensating to random motor and timekeeper variability and thus for maintaining synchronisation with a metronomic pulse (Pressing, 1998; Semjen et al, 1998, Semjen et al., 2000; Vorberg & Wing, 1996). Phase correction models have also been argued to be able to solely account for deliberately introduced fluctuations around a constant mean period (see Repp, 2001). Interestingly, experience of participating in tapping studies, but not musical expertise, has shown to reduce the phase correction response (Repp 2010b). This has been viewed as improved performance in that reacting less forcefully and spreading out phase correction over a few taps decreases the variability of asynchronies and inter-tap intervals. In general the length of the inter onset interval positively affects the size of the phase correction response (Repp 2008c; Repp 2011c)[45]. Furthermore, it is worth noting that participants can anticipate a phase correction to expected perturbations and further reduce timing errors (Repp and Moseley, 2012; see Van Steen & Keller, 2013).

Phase correction seems to be a predominantly automatic process. It operates by and large without awareness in participants and is equally effective within and beyond perceptual thresholds for detecting perturbations and/or asynchonies (Repp 2000; 2001; Repp & Penel, 2002). This highlights that the internal processes are directly coupled to the sensory information without any mediation by higher cognitive function (Repp, 2001). As such this perception-action coupling (Kelso & Kay, 1987) is part of a wider phenomenon of sub-perceptual guidance and evidence against a necessary conscious perceptual mediation between sensory input and action (Neumann, 1990). Repp and Keller (2004) further show that phase correction is unaffected by a diminution of attentional resources (e.g. a concurrent mental arithmetic task) and hence, unlike period correction, does not appear to require attention. Intention also, only has partial effect. Participants instructed not to react to perturbations in the stimulus phase show a reduction in the phase correction response but it cannot be entirely suppressed in the tap immediately after the perturbation (Repp 2002a; 2002c; Repp & Keller 2004). In perturbations below perceptual thresholds there is no effect of intention. Indeed this experimental finding may be partially indicative of why unintentional forms of auditory, locomotive, and visual entrainment occur in naturalistic settings (Schmidt and O'Brien, 1997; Richarson et al, 2007; Coey et al, 2011; Demos et al, 2012; Oullier et al, 2008; Van Ulzen et al, 2010; Sofianidis et al, 2012) and also partially indicative perhaps of why we find it almost impossible sometimes to avoid locking-in to another's musical beat (Lucas et al, 2011). It seems, effectively, that phase correction is only subject to voluntary control if we are aware of the perturbation.

Period correction only seems to be implicated in synchonisation tasks if the timing of the external stimulus sequence is modified (Repp, 2001). Naturally phase correction alone become insufficient for maintaining synchronization when systematic tempo fluctuations exceed a certain threshold in relation to other parameters, most notably the base tempo (e.g., Takano and Miyake, 2007). Period correction adjusts the period of the timekeeper that drives the motor activity with the resulting new period persisting until a further correction is applied (Repp, 2001b). The clearest examples of period correction in operation are therefore to be found in tempo changes to the pacing signal. Michon (1967) first introduced sudden tempo changes (of 8% or more) to isochronous sequences showing that participants were able to synchronise (to within pre-alteration variability levels) after 4 to 5 taps. Michon also showed that an initial overshoot in the subsequent inter-tap-intervals (ITIs) was followed by an individually varying gradual approximation. His initial mathematical analysis that period correction is *solely* responsible for tracking pulse changes has subsequently been renounced in favour of combined models incorporating both phase and period corrections. The initial overshoot in correction makes perfect sense if we suppose that both processes independently modify the timing of the next action based on respective percentages of the asynchrony (Repp and Keller, 2008) or the difference between the preceding inter-onset interval and the preceding timekeeper interval (Hary and Moore, 1985, 1987). Two-process error correction models that account for short-term phase

---

[45] Interestingly, however, given the development trajectory of entrainment capabilities the age of the participant shows no effects (Turgeon et al, 2009).

correction and longer-lasting period correction within a single model (Mates, 1994a, b; Vorberg and Schulze, 2002; Vorberg and Wing, 1996; Semjen et al., 1998; Schulze et al., 2005) have variously estimated phase and period correction values between 0.2 and 0.8 depending on variabilities in tempo, task, and analytical approach (Repp and Keller, 2004; Fairhurst et al., 2012; Repp et al., 2012).

The perceptual liminality of the perturbation appears to be a crucial factor. Hary & Moore (1985, 1987a) were the first to employ subliminal (-1.3% (10ms)) changes to a baseline IOI of 700ms. In contrast to supraliminal alterations, results for subliminal (10/20ms) step changes showed a rapid adaptation *without* an initial overshoot. Hary & Moore initially proposed a mixed phase resetting model for this finding in which period correction occurs slowly in combination with ‘mixed resetting’. This has subsequently been shown to be the formal equivalent to a phase correction model in which the current timekeeper interval is constantly adjusted by a proportion of the most recent asynchrony (Schulze, 1992) and has become widely known as a linear phase correction model (Mates, 1994; Vorberg & Wing, 1996; Pressing, 1998; Semjen et al., 1998, 2000; Repp, 2000; 2001) (see Schulze & Vorberg, 2001 for an overview). It is important to highlight that these, however, do not account for the initial ITI overshoot in supraliminal corrections and predict, therefore, that a period correction mechanism is additionally operational in supraliminal conditions. Thaut, et al. (1998) supported this hypothesis in an experiment in which they used cosine-wave modulated metronome sequences employing both subliminal and supraliminal step changes. As well as confirming the presence and absence of an initial correction overshoot in supraliminal and subliminal corrections respectively, they also found that asychronies returned rapidly to the pre-step change level with supraliminal step changes, however, only very slowly after subliminal step changes. These results, showing that internal period correction increases with step change magnitude as a function of awareness, have been subsequently corroborated (Repp 2001b; Repp and Keller, 2004) and supported by a study exploring the potential neuroscientific correlates of this distinction. Stephan and colleagues (2002) showed that while ventral prefrontal cortical areas are engaged in subconsciousness motor adaptations to auditory stimulus, conscious motor adaptations includes *additional* dorsolateral prefrontal cortex involvement.

Period correction is effectively a more challenging cognitive task. It is largely under cognitive control, requires attentional resources, and relies on the conscious perception of a tempo change in the pacing sequence. This was shown clearly in an important study by Repp and Keller (2004) who found that period correction is dependent upon variables of intention, attentional load and awareness, whereas phase correction was only affected partially by intention in supraliminal conditions (see above). Thus, they argue that phase correction and period correction seem to represent independent processes of largely automatic action control and of intentional cognitive control, respectively (Repp and Keller, 2004). They hypothesise that their findings reflect the notion that “period correction is based on a more complex form of sensory evidence – namely on a difference between intervals (“relative period”) – than is phase correction, which is triggered by a difference between time points (or relative phase)” (p. 517). In effect they are highlighting the distinction that period correction can be characterised as an interval comparison - a second-order difference - whereas phase correction results from a simple phase discrepancy - a first-order difference. The latter they posit “even inanimate dynamic systems can perform” (p.517). Another way of explaining this idea, in the context of the effects of intention only, is given by Repp (2001b) who hypothesized that the differences were due to “period correction requiring memory for at least one preceding event and as such greater computational complexity calling for greater neural resources, thus making the process more extensive in brain space and in time, and hence more accessible to higher-level cognitive processes” (p. 310-311). Extending on this in a later paper Repp (2004, p. 76) states that “it is likely that period correction is a specifically human ability [and] is a manifestation of the more general human ability to set the tempo of a rhythmic activity at will.”

Of course, in real musical settings additional social and interactional capabilities are operational in achieving temporal and physical coordination (see below). Nevertheless we can safely assume that these observations of two correction processes reflect part of the basic temporal framework for real-time interpersonal musical behaviours across cultures. It seems entirely plausible that the relative reliance on these foundational mechanisms in music increases, revealed perhaps by a ‘stricter’ adherence to pulse, with the challenges of larger group size and lack of social sympathy, familiarity and experience (see below). Phase correction mechanisms can be supposed to be operational in all activities involving future-directed attending where expectations are constantly updated based upon asynchronies between attentional pulses and stimulus events (see Bispham 2006). In contrast, period correction is almost by definition functional, specifically, within the framework of a sustained musical pulse. It is this latter process that arguably ties together the recent past, the psychological present, and our immediate expectancies into a longer phenomenological experience. As such, I argue that it forms part of a foundation for the specificities and particular efficacies of music (see below).

Sensorimotor synchronization in the animal kingdom

It is with the above discussed literature, and ultimately also against the embeddedness of human musical pulse in movement and sociality (see below), that examples of animal entrainment/beat matching need to compared if we are understand potential cross-species analogies and/or homologies. Alternately, proposals as to which animals exhibit “entrainment” and the putative relevancies to human capacity and musical capabilities risk being misunderstood and could be ultimately undermined by varying terminological boundaries. Wilson and Cook (2016), for example, suggest that entrainment may be more widespread than we currently have evidence for and include discussions of ‘neurological plausibility’. Nonlinear oscillations and neural entrainment are indeed an inherent principle of brain functioning across species and intrinsic to the physics of perceiving, attending and responding to auditory stimuli (Large and Synder, 2009). They are very likely embedded preconditions of musical pulse and rhythm but also inherently insufficient (see Bispham 2003; 2006; Fitch, 2012). Fitch describes this as the “paradox of rhythm” in that ‘‘periodicity and entrainment seem to be among the most basic features of living things, yet the human ability (and proclivity) to entrain our motor output to auditory stimuli appears to be very rare'' (Fitch 2012, p.28). To avoid confusion, the evidence of beat-matching in animals is, therefore, discussed below in terms of sensorimotor synchronization rather than entrainment more broadly.

Snowball the Cockatoo and (Vocal-learning) Friends

In 2006 Patel first suggested that vocal learning - the ability to accurately mimic complex, learned sounds - may have provided a necessary, but not entirely sufficient, precondition for the capacity for 'beat perception and synchronization' (BPS). He built his 'vocal learning and rhythmic synchronization' hypothesis on the observations that BPS depends upon neurological circuitry used in complex vocal learning. They both involve particular connections between the auditory and motor systems in both birds and humans (Doupe et al, 2005; Patel et al, 2005) and both seem to share overlapping neural substrates in the basal ganglia and supplementary motor areas (Jarvis, 2004; 2007; 2009; Chen et al, 2006; Grahn and Brett, 2007). Interestingly, the evidence suggests that vocal learning evolved independently in three distantly related groups of birds (parrots, songbirds, and hummingbirds) and involved similar sets of brain nuclei. Fitch and Meitchen (2013) therefore argue that this represents a 'deep homology' – convergent evolution of similar biological mechanisms due presumably to genetic constraints on how a capacity can parsimoniously be manifest. An additional point of potential connection to human vocal learning and the motor control for human speech (and presumably music) is that the famous FoxP2 gene is also expressed in avian brain regions important for motor control of learned song (see Jarvis 2013; Chen et al 2013). Patel's hypothesis accommodated the fact that humans appear currently to be unique among primates in exhibiting complex vocal learning (Egnor & Hauser, 2004) and predicted that only species who exhibited vocal learning (e.g. parrots, dolphins, and seals) could potentially be capable of BPS (sensorimotor synchronisation). Accordingly the majority of species that don't exhibit vocal learning, even human pets (e.g. dogs and cats) were expected not to be capable of BPS.

Subsequent internet fame for Snowball the "dancing cockatoo" seemed to support the hypothesis, as did an analysed search of YouTube videos. Keywords "dance" and selected animals brought up 33 videos in which exclusively vocal mimicking animals (14 species of parrot and one elephant [see Stoeger & Manger, 2014]) showed evidence of entrainment (Schachner et al, 2009). Of course this approach was scientifically problematic and subject to a myriad of potential selection biases (see Wilson and Cook, 2016). Furthermore it was unclear whether the animals were responding to human visual cues or could synchronise to a range of tempi and adjust to any changes. Snowball and 8 budgerigars have been studied under more experimental conditions. In Patel et al (2009) Snowball was tested to excerpts of "Everybody" by the Backstreet Boys (a song he was very familiar with). The original (tempo = 108.7 bpm) was manipulated to 10 additional tempi with maintained pitch (+-2.5%, +-5%, +-10%, +-15%, and +-20%. – [97.8 - 130.4 bpm]). His performance was perhaps not as impressive as the initial videos suggested. He did not consistently synchronise to the beat. However, he did exhibit statistically significant bouts of entrainment in 58% of trials in which he showed sustained "dancing" and with all but the two slowest tempi and more frequently with the faster tempi. He showed marked preferential areas for synchronization. The apex of headbobs occurred typically after the beat at an average phase angle of 3.94 (not significantly different from 0) and therefore didn't exhibit the negative mean asynchrony we humans typically exhibit, but did suggest some anticipatory mechanism in that it was too fast to be a reaction to the stimulus. More impressive perhaps were a group of eight budgerigars (*Melopsittacus undulatus*) who have been shown, after relatively little training, to "tap" 6 beats to acoustic and visual pulses at a range of tempi (Hasegawa et al, 2009). Curiously they found slow frequencies (1200 – 1800ms) easier in terms of learning the task but were most accurate at faster rates (400 – 650ms). Despite reports to the contrary in Repp and Su (2013) they also differed significantly from Snowball in that they typically moved ahead of the beat. As yet, to my knowledge, this is the only study to report experimentally on many individuals of one species synchronizing movements to a beat. As such, budgerigars would definitely appear to merit further research.

Ronan the Sea Lion

A challenge to the vocal learning hypothesis comes from studies on a Californian Sea Lion (*Zalophus californianus*) – 'Ronan' – a species that has not been documented to exhibit vocal mimicry. In impressive bouts of synchronization Ronan entrained to five novel tempi of frequency-modulated stimuli (96 bpm, 108 bpm, 88 bpm, 132 bpm, and 72 bpm). This notably included head-bobbing to experimental complex sine tones and also to the same Backstreet Boys song used in Snowball's trials (notably a song with a strong, regular, and markedly overt drumbass). His synchronization was unperturbed by a missing beat in the stimulus but, unlike human tappers, was generally lining up the end of his movement trajectory slightly behind the beat (Cook et al, 2013). For reasons that remain unclear, Ronan was generally farther behind the beat for faster stimuli and farther ahead of the beat for slower stimuli and interestingly showed improvements with increasing familiarity with a stimulus. A subsequent set of experiments showed some even more notable findings in that Ronan tracked temporal phase and period perturbations to an artificial click track, including in novel tempo conditions (Rouse, 2016). The authors argue that Ronan's beat keeping across all tempi (61.818 to 125.925 bpm) and adaptations (±25% of the IOI) impressively fit dynamic models of co-oscillation. Crucially for our purposes, the analysis showed changes in phase and tempo were matched through both phase and period adaptation. As such it is the only documented case of a non-human animal exhibiting period correction in sensorimotor synchronization and challenges my and Bruno Repp's suggestion that period correction is unique in humans (Bispham, 2006; Repp 2004). However, there were also two important differences from human studies. First of all, phase coupling varied with perturbation magnitude and was particularly increased for presumably noticeable perturbations beyond 8 %. Secondly, the relative period adaptation was far lower than in human studies. Ronan's period coupling values did not exceed 0.2 whereas in humans it is typically varying values between 0.3 and 0.8 (Loehr et al., 2011). Furthermore, rather than a single sudden shift, the tempo was shown to change continuously following the shift onset. Given that description it would perhaps be interesting to explore if these corrections could have also been explained within an information-processing linear phase-correction model (e.g. Mates, 1994; see above).

The origin of Ronan's capabilities, which are not reflected in any naturalistic behaviour, remains a bit of a mystery. Related pinniped species have been shown to possess considerable vocal flexibility (Sanvito, Schusterman & Reichmuth, 2007; Galimberti & Miller, 2007; Schusterman, 2008). However, only one individual harbour seal, raised in an artificial environment, has been shown to exhibit vocal mimicry (Reichmuth & Casey, 2014). Different species also diverged 23+ million years ago so it seems unlikely that there is any common genetic predisposition towards vocal learning in pinnipeds. Another potential candidate that may provide insight into some of the underlying mechanism is working memory. This shares with sensorimotor synchronization the necessity of keeping in mind the immediate past, maintaining it in the present and constructing expectations of the near future. Furthermore, it would correlate with the significance of the prefrontal cortex in both period correction adaptation and in working memory tasks (Stephan et al, 2002; D'Esposito and Postle, 2015). Notably, California sea lions have been shown to possess good visual and auditory working memory (Schusterman and Kastak, 2002) and pinnipeds' auditory working memory even exceeds, in certain conditions, that of non-human primates (Fritz et al., 2005, Scott et al., 2012; see Ravigniani et al 2016).

Synchronous chorusing and coalition signaling

The avian and pinniped studies discussed were almost entirely constructed, artificial and lacking in naturalistic correlates. Examples of naturally occurring group periodic synchronisation and coordinated vocal displays, while lacking comparatively in experimental vigour, do show some consistency in terms of context and function. These have, therefore, prompted speculations regarding their homological relevance to the evolution of musical pulse in human. Merker (2000) argues, that as periodic group synchronicity is restricted to male sexual display in the animal kingdom, it is likely to have played a similar role at some point in recent (from 6-7million years ago) human evolution and supports this view with reference to female exogamy in humans and chimpanzees (Ember, 1978). Documented cases include bioluminescent flashing in fireflies, claw waving in fiddler crabs, and 'chorusing' in Neotropical katydid and some frog species (see Greenfield, 1994). In an essentially competitive context, synchronicity is achieved either as an indirect outcome of all participants desiring to signal first; as a cooperative effort to maximise output and hence attract females to their group; or as an individual means of avoiding predators. In all cases the synchrony is based on either sight or sound. It is, however, achieved either by advanced signalling or through phase correction mechanisms, and occurs only within a narrow temporal range corresponding closely to each individuals 'eigenfrequency' for the particular action. Additionally the 'pulse' does not act in any way as a reference point for any other action. Thus this view does not sufficiently account for a lack of contextual, physiological or behavioural correlations between the examples given and synchronised behaviours in humans[46]

Musical Pulse as Movement

"Musicality, as the activity of a unitary "self" or "individuality" expressing motives to move, is organised in one connected time frame. It cannot be synthesised or accreted in the mind from separately anarchic processes, but must have one coherent and orderly rhythmic foundation. It expresses an integral stream of events in the whole brain, which conduct separate body parts to targets in a real or imagined space-time of experience, synchronising moves so the effects of separate actions can balance one another and form anticipated sequences and coincidences in space and time, as nearly faultlessly as possible. The gracefulness of all we do depends on it" (Trevarthen 1999/2000, p. 160).

Regardless of how we interpret Ronan's surprisingly impressive capabilities, it seems unlikely that the behaviour has any great psychophysiological significance beyond the practicalities of the task. Crucially in humans, however, a sensorimotor synchronization task, or more broadly entraining to a sustained pulse, connects inextricably with the fundamentals of motor functioning and aspects of our capacity for culture. It is essential to an evolutionary perspective on music to appreciate that the capacities inherent in a musical pulse are inextricably connected and shared with those that facilitate movement and an appreciation of volitional and non-volitional action in time. A musical pulse is, in essence, an extended form of dynamic action. The following sections will explore how pulse and beat perception are embedded with action simulation and movement from early developmental stages in humans. I will also show that pulse and beat perception share considerably with more fundamental mechanisms that afford smooth motion. Therefore a key question I will address below is whether, and to what extent, evolutionary changes to our capacity to plan complex action sequences can account for the specificities of a musical pulse.

---

[46] A related perspective on the human ability to coordinate and synchronise actions is Hagen and Byrant's hypothesis (2003; Hagen & Hammerstein, 2007; see also Fessler & Holbrook, 2016) that music and dance evolved as a coalition signalling system that could reliably communicate coalition quality, thus permitting meaningful cooperative relationships between groups. They argue that this capability may have evolved from combined territorial defence signals that are common in many social species, including chimpanzees, gibbons, lions, hyenas and wolves (e.g. Herbinger, Boesch, & Rothe 2001; Mitani, 1990; Heinsohn, 1997; Hofer & East, 1993; Harrington & Mech; 1983) and have become increasingly coordinated in humans as a signal of group cohesiveness. While music is seemingly able to fulfil this function it seems unlikely that this represents the primary adaptive functionality of rhythm and entrainment. Crucially the animals in the examples given do not exhibit efforts to genuinely coordinate their calls beyond being engaged over a similar time span and offering and reflecting encouragement to participate. It also fails to account for the generic complexities of music, the full scope of entrainment in humans (ranging from group synchronicity of physicalised action to more subtle manifestations in, for example, mother-infant interactions); the underlying psychological mechanisms being vastly different; and/or its most prevalent contextualisation in affective interpersonal interaction.

Primarily, music fully enacted *is* movement both in the production and perception of action in time. It is born of movement, it moves us, and compels us to move. It is essentially and universally a form of dance (Nettl, 1984, 2000; Brown, 1991). More precisely perhaps we should state that the culturally specific terminological boundary between music and dance reflects two different manifestations or individually incomplete parts of what is essentially one and the same thing. Ancient philosophers already noted that "everything flows" (Heraclites; see Brandner, 2012; Todd & Lee, 2015), described speaking and singing as gestural motion (Aristoxenus c300BC), and music as sound *and* movement (Quintilianus c 300AD). More recently researchers have presented further intuited descriptions of music perception being, first and foremost, audible human motion (e.g. Truslit, 1938; Gabrielsson, 1973; Repp, 1993; Shove & Repp, 1995; Das et al., 1999) as well as more mathematical models of, for example, structural variations in musical tempo being correlates of motion (e.g. patterns of locomotion) (Todd, 1992, 1995; Feldman et al., 1992; Kronman & Sundberg, 1997; Sundberg & Friberg, 1999). Undoubtedly, part of these connections relate to a coupling of perception and action (Gibson, 1966; Prinz, 1990; 2013); the inherently rhythmic nature of any motor action (Trevarthen, 1999); the complex amalgamations of timing, temporal perception, and motor skills (e.g. Ivry & Hazeltine, 1995; Wohlschläger & Koch, 2000; Schöner, 2002; Ivry and Spencer, 2004); and our intuitive and abstracted understandings of the basic physics of motion (Newton, 1687). Notably neuropsychological studies also support the existence of a common mechanism showing shared brain structures across temporal processing and motor output (Keele et al., 1985; Treisman et al., 1992; Coull and Nobre, 2008), as well as across sequencing, action and music (Janata & Grafton), and time perception and motor timing (see Schobotz, 2000; Avanzino et al., 2016). It is critical to note therefore that, broadly speaking, the strong associations with movement are not *solely* based on features that are unique to music.

Action simulation

Temporal and interpersonal coordination in any rhythmic action and joint action is clearly also supported by complex pre-motor planning in production and an understanding of the timing of motoric actions as well as the corresponding social motivations and goals of those participating, including ourselves (see Keller et al, 2015). Action simulations are well documented sensorimotor neural processes in which those areas engaged in executing an action are reflected in "passive" observations (i.e. in the absence of overt movement (see Gallese et al., 2004; Wilson & Knoblich, 2005; Decety & Grèzes, 2006). These motor simulations form part of our anticipatory mechanisms, generating predictions about the future time course of our and others' actions (Ross et al., 2016). They thus contribute to facilitating precise rhythmic interpersonal coordination in a range of interactions. They can be modelled internally as forward, inverse and joint associations between efferent motor commands and afferent sensory information. Forward models describe the causal relationship between motor commands and their effects on the body and environment whereas inverse models represent transformations from intended action outcomes to the motor commands that produce them (Wolpert et al, 2003; Tian and Pöppel, 2010). Motor simulation is clearly operational in music and dance (Keller, 2008; Cross et al., 2006) with corresponding training effects on visuo-motor (Calvo-Merino et al., 2006) and audio-motor simulation (Lahav et al., 2007) respectively. Studies focusing on music show, for example, increased motor excitability while listening to a piano piece after learning to play it on the piano (D'Ausilio et al., 2006). Broadly speaking it is, however, certainly not unique to music either. Notably many neuroimaging studies also show evidence of motor activation while adult and even pre-verbal infant participants listen to speech (Skipper, et al., 2005; Wilson et al., 2004; Kuhl et al. 2014). More specifically there are suggestions that speech effector muscles show facilitation when listening to speech (Fadiga, et al., 2002)[47]. Finally it is important to note that one can take this further to incorporate more abstract or internalized representations of form in motion. In a full sensorimotor perspective a perceived rhythm is, literally and neurophysiologically, an imagined movement. Todd (1999; see also Todd et al., 1999), for example argues that the same mental processes that generate bodily movement are involved in listening to music (or perhaps perceiving any complex rhythmically organised sound). In particular, he argues for a sensorimotor loop (including the posterior parietal lobe, premotor cortex, cerebrocebellum and basal ganglia) overlapping considerably with the mental structures serving timed motor action - the 'motor circuit'. He supposes that "if the spatiotemporal form of certain sensory stimuli are matched to the dynamics of the motor system, then they may evoke a motion of an internal representation, or motor image, of the corresponding synergetic elements of the musculoskeletal system, even if the musculoskeletal system itself does not move" (p.119).

Beat perception as motor action in the brain

Crucially for our comparative investigation of a musical pulse, a substantial body of evidence suggests that motor system activation in music may not merely be a reflection of sequential planning, action, suppressed movement, or abstraction. Even in absence of any particular shapes and enacted or imagined movements, the embedded connections between a sustained attentional pulse, rhythm perception, and movement seems to go to the very heart of motor and vestibular functioning (Thaut et al., 2015; Todd and Lee, 2015). Neurological studies support a deep-rooted connection between a musical pulse and movement (and also with timing and duration perception [see Merchant et al., 2013). Activations in studies of rhythm and beat perception, using even the most basic representations of a "beat", appear consistently in areas of the brain associated with motor action, preparation and timing. Specifically, perceiving and producing musical rhythms correlates with the supplementary motor area, premotor cortex, cerebellum, and basal ganglia (Bengtsson et al., 2009; Chen et al., 2008a; Grahn and Brett, 2007; Grahn and Rowe, 2013; Kung et al., 2013; Lewis et al., 2004; Marville et al., 2002; Schubotz and Van Cramon, 2001; Ullen et al., 2003). A particularly relevant observation from these FMRI studies is that

---

[47] There are notably transparent parallels between the theories of motor simulation and mirror neurons (see Koelsch, 2012). However, one interesting difference, as noted in Ross et al (2016) is that the mirror neurons associated with the ventral premotor area do not appear to be related directly to the more dorsal premotor areas that are associated with beat perception.

patients with impaired basal ganglia function due to Parkinson's disease (characterized by severe deficiencies in motor functioning) show impairments in a rhythm discrimination task (Grahn & Brett, 2009). This markedly ties in with the importance of the basal ganglia in beat-based versus absolute timing (see above) and discussions of rhythmic facilitation (see below). It seems likely therefore, as I suggested in Bispham (2003), that a subset of the mechanisms involved in musical rhythmic behaviour, temporal processing and dynamic attending are involved, also, in timed, rhythmic and untimed motor behaviour. A critical point, therefore, is that one can posit that the mechanisms that allow us to create a musical pulse are built upon (or perhaps even generative of) those mechanisms that help, architecturally speaking, to structure movement, balance, and the perception of complex actions in time. In the perception of temporally structured events and action we decipher (mostly subconsciously) an internal organizational structure, or manifestation thereof, that affords prediction, efferent and afferent coordination, and a management of smooth and ballistic movement. A musical pulse is, in effect, a particular form of this organizational structure, incorporating period correction and corresponding degrees of volitional control and an awareness of temporal framework.

Pulse and movement in development

Human infants also appear to possess a considerable propensity for rhythm and early connections to movement. A concept of relative processing appears to be primarily salient as infants discriminate between rhythmic patterns based upon the relative durations of events in spite of tempo changes (Trehub, & Thorpe, 1989). Additionally they respond to changes in order (Demany, McKenzie, & Vurpillot, 1977) and rhythmic groupings (i.e. Changes from a 2-2 grouping to 3-1 grouping of the same events) (Chang & Trehub 1977a; Lewkowicz, 2003). However, they also detect changes to absolute durations, showing renewed interest in familiar rhythm at new tempos (Pickens & Bahrick, 1995) and can lose discriminatory abilities if rhythms are presented at much altered tempi (Pickens & Bahrick, 1997). Interestingly, infant rhythm perception capabilities appear to be operational across modalities (Mendelson, 1986). Evidence that metrical categories are in a sense physically embodied in infants comes from a study showing that infants will respond preferentially to metric accentuation of an otherwise ambiguous rhythm if they were previously bounced according to that particular metric interpretation whilst listening in the absence of accentuation (Philips-Silver & Trainor, 2005). In this astutely designed study infants who were bounced to an ambiguous (not accented) drumming rhythm on every second beat subsequently showed greater interest to a duple than triple meter accented version (and vice versa). Crucially simply watching the adult move to the ambiguous rhythm in 2 or 3 did not show the same effect. Therefore, it seems that connections between pulse and movement appear to be established early in development. Extensions to the original study showed parallel results in adults (Philips-Silver and Trainor, 2007) and demonstrated that head movements in particular were necessary for the observed effect (Phillips-Silver and Trainor, 2008). The authors interpreted this as evidence of a primary influence of the vestibular system and supported this in a further study highlighting that the results were mirrored further using galvanic vestibular stimulation (Trainor et al., 2009). However, it has been suggested that the vestibular influence is perhaps not direct (Trainor, 2007; Riggle, 2009; Trainor and Unrau, 2009) and, in general, vestibular models of beat perception in development and maturity (Todd, 1995; Todd and Lee 2015) remain somewhat controversial with most researchers focusing on more ballistic accounts of the connections between musical pulse and movement[48]. Either way, the evidence seems to have securely demonstrated that the connections between pulse and movement arise naturally and are already embedded in early developmental stages.

Rhythmic facilitation in rehabilitation

Perhaps the most insightful area of research highlighting the connections between a musical pulse and movement comes from studies of motor rehabilitations through rhythmic stimulation. Thaut and colleagues (1992) first showed improvements in the gait of healthy individuals with "rhythmic" (an isochronous pulse) cuing. In particular, stride symmetry and electromyography patterns in amplitude variability in muscle contractions across the stride cycle were shown to be markedly improved. Comparing results of the differential use of metronome vs. musical rhythmic-auditory stimulation in healthy subjects showed greater precision with the rhythmically accentuated musical stimulation (Thaut et al., 1997). However, the difference was small indicating that simply the pulse itself was responsible for the majority of the observed effect and gait improvements. Subsequently, similarly positive but even more dramatic results of hemiparetic motor stability were shown in stroke patients (Thaut et al., 1993). Analysis showed immediate positive modifications: The stride lengths of the affected and non-affected sides became more symmetrical; hip joint range of motion of the affected/non-affected sides became more symmetrical; and the centre of mass vertical displacement decreased, indicating improvement in mechanical efficiency (Thaut et al., 1997). Interestingly, the results were significantly superior to those achieved through physiotherapy (see also Thaut et al., 2007; Hayden et al., 2009). Ensuing results with Parkinson's disease patients were even more spectacular still. Dramatic improvements were evidenced immediately in synchronising step patterns to a music-based auditory rhythm and also in sustained results following a 3 week home-based walking program for 30min per day (Thaut et al., 1996). The

---

[48] Todd and Lee (2015) argue that dancing to music which does not have a beat counteracts oscillator theories arguing that moving to music with ametrical rhythms, such as recitative, plainsong or more modern ambient styles, or much music featuring a deep rubato (an expressive manipulation of the beat [e.g. much Romantic music]) is more likely to involve swaying the whole body along with the larger scale motions associated with phrasing. They subsequently describe such swaying (and also chanting) as a form of vestibular self-stimulation and argue that "it is possible to associate such gestural or whole body self-motion perceptions with somatotopic maps within the medial sensorymotor circuits as a kind of "inner movement" in the sense of Truslit's *innere Bewegtheit*" (p.19). They continue to consider that from a neurobiological perspective "such motional percepts may almost literally be created as was predicted by Truslit, i.e. by either indirect associative links of sound shapes to vestibular centres in the body maps or by direct vestibular activation of body maps by sound above the vestibular threshold" (p.19). Although there are some considerable problems in describing some of the examples given as not having a beat it does raise the possibility that the prevalence and levels of adherence to a regular pulse in music correlates with a connection to ballistic or vestibular functioning respectively.

numbers showed roughly a 25 per cent increase in stride velocity, a 12 per cent increase in stride length, and a 10 per cent increase in step cadence. The numbers are impressive. However, it is observing the transformation from one moment to the next that is most striking. It seems to suggest not just a quantitative shift, an improvement in particular data points, but rather an altogether more harmonious whole – a transformed holistic coordination of the whole body. This is a shift, as Mithen (2005, p.151) describes, from "cumbersome, shuffling movements to a quite unexpected degree of muscle control and fluid motion".

These seminal studies invigorated advocacies for the role of music as a neurological therapeutic stimulus (see Thaut, 2015b; Mainke et al, 2016) and, in particular, the view that its structural properties can shape sensory perception linked to motor function (de l'Etoile, 2010; Altenmueller and Schlaug, 2013). The findings of pulsed and music-based rhythmic facilitation have been extended in many further studies on hemi- paretic gait rehabilitation (Ford et al., 2007; Roerdink et al., 2007, 2011; Thaut et al., 2007; Spaulding et al., 2013) and in the beneficial use of RAS for mobility in Parkinson's disease (see deDreu et al., 2012). Crucially, further extensions have also produced positive results in other patient groups with motor disabilities and also in a comprehensive range of different ballistic and smooth motor-actions. Patients with Huntington's disease, spinal cord or traumatic brain injury (Bilney et al., 2005; de l'Etoile, 2008; Adler & Ahlskog, 2000; Hurt et al., 1998); children and adults with cerebral palsy (Peng et al., 2010;Wang et al., 2013; Ladwig et al., 2016); and cerebellar stroke victims (Wright et al., 2016) have all benefited from rhythmic auditory stimulation. The actions and kinaesthetic markers studied have variously included kinaematic parameters of arm movements (Safranek et al., 1982), limb instability (Thaut et al., 2002); reaching movements (Ladwig et al., 2016); circular arm movements (Massie et al, 2009); repetitive cyclic motions for upper extremities; and markers of full body coordination (the latter especially in hemiparetic stroke rehabilitation and children with cerebral palsy) (Luft et al., 2004; McCombe- Waller et al., 2006; Schneider et al., 2007; Altenmueller et al., 2009; Malcolm et al., 2009; Grau-Sanchez et al., 2013; Peng et al., 2010;Wang et al., 2013). Control of speech has also been shown to benefit from rhythmic entrainment with demonstrable improvements in intelligibility, oral motor control, articulation, voice quality, and respiratory strength in patients with Parkinson's and post-stroke aphasia (Wambaugh and Martinez, 2000; Thaut et al., 2001; Lim et al., 2013). Notable, also, is the fact that the success of melodic intonation therapy in aphasia rehabilitation has been partly attributed to the rhythmic component (Stahl et al., 2011). Links to areas of the brain implicated in motor timing are implicit in the selection of tasks, disabilities and conditions studies but were also made explicit in a comparative study of walking in patients with differing brain lesions (Kobinata et al., 2016). Rhythmic auditory stimulation resulted in improved velocity and stride length in patients with cerebellum, pons and medulla, and thalamus lesions – all areas heavily involved in motor planning and timing. Gait parameters in the putamen and corona lesion patients were, however, unaffected.

Despite the increasingly overwhelming functional evidence, the neuroanatomical and neurophysiological basis for the improved kinematic parameters described above remain yet to be fully understood. It seems likely, as suggested by Thaut and colleagues (1999), that the enhanced time information inherent in the regularity of the temporal structure of rhythm (pulse) serves as an optimisation function for motor planning, programming, and execution. It serves, effectively as a reference point in the continuity of action in time. Clearly a subset of the mechanisms involved in beat induction and dynamic attending are involved, also, in motor behaviour. This is a crucial point highlighting, similarly to earlier sections on our experience of time and temporal perception, that our capacity to engage with a musical pulse did not evolve independently. Rather, it is, in a very real sense, a particular extension of our fundamental capacities for perceiving action in time, and for movement. It is interesting to note that the benefit level seems to increase with the severity of motor dysfunction, thus suggesting some kind of compensatory operation and that entraining to rhythmic stimulus provides access to *or* simulates periodic oscillations which are crucial to motor control. However, we do not know which of these latter options is correct - whether the compensation occurs directly or indirectly. Does auditory entrainment to a pulse provide access to periodic internal mechanisms that are not functioning naturally due to the respective motor disorders or does it provide external reference points as a distinct way of achieving similar results but using different pathways to those operational in motor functioning in normal subjects? Either way it does seem, as proposed by Thaut (2005, p. 58), entirely plausible from the results of these studies that "the basic neural network underlying isochronous pulse synchronization consists mainly of composite motor and auditory areas with no clearly designated functionally separate area for synchronization".

Motor skills and beat-based timing in human evolution

The above discussion on musical pulse as movement and, especially, the rhythmic rehabilitation studies raise some intriguing and critical questions concerning the evolution of beat-based timing and sensorimotor synchronization in humans. We have already seen some well-supported arguments for much tighter auditory-motor connections in humans than in other primates (Patel & Iverson, 2014; see above). However, as the connections between beat-based timing and motor control seem to be so tightly interconnected, it remains somewhat puzzling that humans are one of the only species to show beat-based timing and entrainment. One possibility I suggested previously (Bispham, 2003) is that increased cognitive cross-domain flexibility in the hominin lineage (Mithen, 1996; see Section II introduction) allowed dynamic oscillatory mechanisms that are fundamental to motor action in all species to be abstracted, made conscious and replicated in pulse-based and musical contexts. This perspective markedly does not assume that the mechanisms that serve bodily movement in humans are fundamentally different from those affording kinaesthetic movement in other animals and, in particular, that it has remained similar across primate species. However, we should perhaps also consider the possibility that beat-based timing and our sense of pulse emerged exaptively in conjunction with evolutionary changes to motoric capabilities since the divergence from our last common ancestor.

One of the most notable evolutionary changes has been a move to bipedalism (see Harcourt-Smith, 2015; Klein, 1999; Lewin, 1999). Trevarthen (1999) has eloquently argued for the importance of bipedalism for rhythmic capabilities in humans. He states that "while walking, we freely turn and twist, glance with eyes jumping to left and right, extend waving limbs, make intricate gestures of the hands, talk, all in coordinated phrases of flowing rhythm' and that 'this moving has a multiplicity of semi-independent impulses, a potentiality for multi-rhythmic coordination that is surely richer than any other species possesses" (p. 171). This may well be true. However, while this could easily account for increasingly complex possibilities and differences in the rhythmic manifestations of action in humans, it cannot, in itself, directly explain the emergence of an attentional mechanism. Primates are well known to be exceptional among mammals in terms of the number, extension, and complexity of motor-related areas observed in the frontal lobe. Mendoza and Merchant (2014) argue that this is probably related to the wide behavioural flexibility that primates display and that "the increase in the complexity of the motor system has contributed to the emergence of new abilities observed in human and nonhuman primates, including the recognition and imitation of the actions of others, speech perception and production, and the execution and appreciation of the rhythmic structure of music" (p. 73). As humans we have become, in particular, further exceptional in terms of the intricacies of bodily gesture (see Kendon, 2004) and have acquired substantial improvements to fine vocal and manual control as part of the evolutionary drives towards increased communicative and technological ability (Lieberman & Crelin, 1972; Lieberman, 1984; Fitch 2000; Wilson, 1998). Pre-motor planning, evidenced both in complex vocalisation and manual productivities is also exceptional in humans (see Morley 2015), as is a greater voluntary control over our actions and responsibilities (see Haggard, 2008).

An intriguing possibility is that the impressive control of motor action in humans did not simply evolve qualitatively over time, but that it is reflected in a substantial architectural shift in the foundational mechanisms that coordinate movement. Donald (1991), for example, postulates a major transition in hominin cognitive evolution around -1.5 M years (coinciding with transition periods leading to the speciation of *Homo erectus*) in the form of a fundamentally new supramodal motor-modelling capacity – mimesis[49]. Boyd (2018) argues that subsequent hominin evolution research, in particular recent fossil and genetic evidence, more likely dates the proposed pre-linguistic mimetic mind to *Homo erectus* or *Homo heidelbergensis*, about 1.5–0.5 million years ago. In principle, however, the theory remains pertinent and highly influential in discourse into the origins of human sociality and language (e.g. Donald, 2017; Tomlinson, 2018). A crucial property of mimesis, in stark contrast to the supposedly exclusively episodic memory in Apes, is the voluntary retrieval of action-models that allows us to rehearse and refine movements and to create models of reality that are not necessarily tied to the physical environment. This early motor adapation could perhaps have encompassed an additional form of relative timing and/or a capacity to implicitly employ sustained 'pulsed' attentional reference points in modelling action. Thus beat-based timing may have emerged exaptively from this[50]. Increased voluntary control of action may also have played an important role in the evolutionary history of period correction mechanisms in sensorimotor synchronization. Although the issue is philosophically complex and hard to pinpoint experimentally, Haggard (2008) explains that we can perhaps view voluntary action as one end of a continuum that has simple reflexes at the other end. In voluntary action, the occurrence, timing and form are not directly (or perhaps only very indirectly) determined by any identifiable external stimulus. It thus demonstrates a "freedom from immediacy" (Shadlen & Gold, 2004). In terms of a musical pulse this ties in especially well with period correction mechanisms. These, as described above and in contrast to those operational in phase correction, are more distinct from immediacy (they occur at a longer lag), cognitively complex, are correlated with an awareness of the pulsed framework, and are subject to attention and volitional control (Repp & Keller, 2004). We can therefore also not discount the possibility that the most defining feature of a specifically musical pulse, and its capacity to act as a phenomenological bridge between the immediate past, present and future expectation, emerged in human evolution in conjunction with increases in the central executive (Norman and Shallice, 1996), sustained attention (Sarter et al., 2001), working memory (Baddeley, 2003), top-down attention (Baluch and Itti, 2011) and the volitional control of action.

Temporal Structuring in Communication

As well as being embedded in motor action, a musical pulse is also, in its essence, inextricably social. Broadly speaking synchronizing and entraining with others in a myriad of musical and non-musical forms seems to produce and/or be motivated by interpersonal interconnectedness, cooperation, and prosocialities. These psychosocial connections and drives will be discussed in Chapter 4. Here I will continue to focus more on the structural and coordinative aspects of interpersonal interaction. However, this divide is rather artificial, not least in that intersubjective understanding provides considerable coordinative insight. Music is, broadly contextualized, both a particular form of joint action and communicative interaction. In the following sections I will briefly review mechanisms by which interpersonal coordination is achieved in rhythmic interaction. By no means are all of these unique to music or even directly related to specific features of a musical pulse. As will be discussed and analysed, many are broadly enacted in a full range of human and non-human interactive contexts. I will argue, however, that, although some forms of temporal structuring and dynamic expectancies can be discerned, a

---

[49] Crucially Donald (1991) describes the considerable sociocultural implications of mimesis and argues that it is *the* most basic human thought-skill, and can account for the documented achievements of Homo erectus. It is also critical to note that his argument posits that mimetic skill is dissociable from linguistic capacities and retains its own cultural and communicative usefulness (e.g. see Lewis, 2014). It is, in his view, however an absolutely essential preadaptation for the later evolution of language.

[50] Donald highlights the connection to pulse by arguing that the human propensity for rhythm provides the clearest evidence for a unified central "controller" for body mimesis. In musical rhythm motion patterns are abstracted into a perceptual event and transferred almost instinctively from one skeletomuscular system to another. However, it is according to Donald the Gestalt – the overall pattern – that is primary. Thus, it is crucial to note that mimesis describes more general capacities and need not necessarily incorporate all aspects, or the specific features, of a musical pulse.

sustained attentional musical pulse is not subliminally operational in non-musical human interactions (e.g. mother-infant interactions, linguistic rhythm, or interpersonal turn-taking).

Keller and colleagues (2015) offer a music-focused review and discussion of research into the cognitive-motor processes of rhythmic interpersonal coordination that enable individuals to represent joint action goals and to anticipate, attend and adapt to others' actions in real time. They argue for an inclusive model that includes social factors, shared goals and strategies, knowledge, and mechanisms of anticipation, adapation, and attention. The latter three are discussed primarily in terms of the sensorimotor synchronisation skills discussed above and seem to be implied to relate to all forms of rhythmic joint action. As this chapter testifies I am not convinced that they are necessarily *fully* operational in non-musical contexts. Regardless, however, it is clear that all the parts of the model are active in musical ensembles and most (if perhaps not all) are much more broadly operation in joint action. Indeed we can posit a general tendency to rely to varying degrees on those architectural features that seem defining of a musical pulse – the mechanisms involved in periodic sensorimotor synchronisation, its sustained nature and period correction – depending inversely on the relative level of access to social cues and familiarities. This accords well with a widespread increase in levels of synchronisation and a strict adherence to periodicity with increases in ensemble size (i.e where a reliance on individual social cues becomes progressively less valuable and the necessity for clear architectural guideline is increased). It also resonates with the common finding across cultures and styles of *decreased* levels of precise synchrony and a less isochronous pulse in established partnerships or with greater familiarity with a piece or musical style (e.g. Doffman, 2017; League, 2017). On first consideration, this may seem somewhat paradoxical as we often seem to assume (falsely) that higher levels of precision represent an increasingly positive attribute. However, music need not be precise. It needs to be coordinated and harmonised over a sustained period of time. Thus a high precision of synchronisation and a high stability of pulse can sometimes be indicative of a predominant reliance on the structural mechanisms of sensorimotor synchronisation where social cues and points of anticipation based on stylistic familiarities are less accessible. Equally, for example, two musicians in a well-rehearsed and/or established duetting partnership can easily attune to each other's actions, individual characteristics and creative goals. They can stretch an underlying pulse and accommodate each other almost at will. Alternately a mutually understood internalised pulse may remain psychologically stable within a group of, for instance, highly experienced jazz musicians (e.g. Doffman, 2017) even though some, or even all of the group are playing off and around the beat most of the time. The ability to stray from a strict pulse and follow action cues, stylistic features, expressive microtimings and/or interpretations of shared goals can thus be enabled in music precisely because the attentional foundations of musical group coordination are very secure.

In short we need to appreciate that temporal coordination in musical interaction is achieved in large part through aspects of our capacity for culture, our action resonances with others, empathic understandings of common history, experience and shared goals, and broad manifestations of intersubjective attunement. These capacities can be accessed in any temporally structured joint action and/or interpersonal interaction. The crucial comparative question then for our current concerns is whether the underlying and psychological coordinative structure that would appear to be indicative of a *musical* pulse – the sustained predominantly periodic pulse and its corresponding mechanisms for continued correction – is present and/or operational in other primate and human communicative interactions.

Duetting in nature

Vocal 'duetting' is found in many species. In particular, over 400 bird species pairs, representing 40% of bird families, combine to produce coordinated duets and observations in primates span all the major clades. The temporal coordination in avian duets is highly diverse and variable. As Mann (2009) reviews: "duets vary in form from loosely overlapping songs to highly coordinated duets where paired birds both adjust the timing and type of phrases they sing to fit those of their partner over the course of the duet" (p. 67). Duet coordination therefore signals how attentive an individual is to its partner and has also been analysed as a form of turn-taking (in starlings in particular, see Henry et al. 2015). In some species two birds have been reported to coordinate alternating calls with such precision that it is impossible to tell that two birds are involved from the auditory signal alone (e.g., plain wren; Mann et al., 2003). Duetting is most common in the tropics and is thought to relate to the fact that birds there frequently hold year-round territories and use the duets as a form of territorial advertisement (Farabaugh, 1982). This, in turn, is associated with birds that form long-term monogamous pair bonds. Another ffeature that may be associable with duetting is sexual monomorphism (similarity between the sexes). Whether or not any of the naturally occurring duet coordination is achieved through truly analogous correction mechanisms to those enabling a musical pulse or, for example, by synchronously commencing fixed action patterns, remains unclear. The connections to bonding and relationship is, however, clearly an interesting analogy and one we also find in more closely related primate species who duet.

De Waal (1988: 202-203) first described pant-hooting of captive bonobos in a way that caught the attention of researchers interested in comparative issues of musical capabilities. He states that 'during choruses, staccato hooting of different individuals is almost perfectly synchronised so that one individual acts as the 'echo' of another, or emits calls at the same moments as another. The calls are given in a steady rhythm of about two per second'. However, this was merely an observational account and may, therefore, have been affected by a degree of human projection. A later study on bonobo distance calls only reported alternate and not simultaneous calling (Hohmann & Fruth, 1994). It was then Geissmann (2000) who explicitly argued for the relevance of primate duetting to the evolution of music. He described pant-hooting by gibbons (*Hylobates* spp.) who "produce loud song bouts that are mostly exhibited by mated and monogamous pairs. Typically, mates combine their partly sex-specific repertoire in relatively rigid, precisely timed, and complex vocal interaction to produce well-patterned duets." (p. 103). As in many avian species these are generally thought to be involved in territorial advertisement and strengthening of pair bonds. The putative rationale that this behaviour advertises pair strength to potential competitors, as well as the hypothesis that it increases the cost of partner desertion is supported by the observation that duetting becomes increasingly well-coordinated over time and by one case documenting a certain amount of adjustment in the duet structure with a partner exchange (Geissmann & Ordeldinger, 1999)[51]. The mechanisms used to achieve coordination between individuals are clearly complex and of considerable interest to studies of interaction and sociality. Observations would seem to suggest strong correlations between the rate of hoots, levels of kinesthetic movement and degrees of emotional excitement in both animals. As such, it could be argued that the pairs are interactively affecting each other's emotional and physiological state within a context of social bonding and that this does mark it as being particularly relevant and partially analogous to a broader view on musical behaviour in humans (see Bispham 2006). However, crucially, there appears to be no direct evidence for the use of an overarching temporal structure. It is unlikely, therefore, that any other of the discussed behaviours represents or indicates a full analogy to a musical pulse in humans.

Complex temporal structuring in the termed forms of 'duetting', 'turn-taking', 'conversational form' and/or 'chorusing' has since and variously been described in a full range of primate species (Mendez-Cardenas & Zimmermann, 2009; Takahashi et al., 2013; Müller & Anzenberger, 2002; Lemasson et al. 2011). Interestingly, interactional rules appear to be learnt through parental instruction during ontogeny. Marmoset (*Callithrix jacchus*) parents are understood to play a direct role in guiding the development of turn-taking by providing feedback to their offspring during vocal interactions (Chow et al., 2015). However, there appears to be no direct evidence of a wholly analogous sustained pulsed framework being necessary or employed in primate interaction. Federuk et al. (2013), for example, provided some analysis on how the acoustic structure of chimpanzee pant-hooting facilitates chorusing in affliative chimpanzee male from two geographically distinct areas of Uganda. Comparing solo and joint display they found some consistent techniques to enable, prolong, and coordinate 'chorusing'. The occurrence of a 'chorus' was facilitated by regulations to the duration of pant-hoot phases based on others' responses. Additionally, once another had joined in the call, "the initiator seemed to adjust the duration of call phases such as the climax to prolong the display and to synchronise the call with the partner's" (p. 1787). Although these are undoubtedly important and fascinating findings of social coordination, they are notably, and in contrast to the mechanisms involved in sensorimotor synchronisation in a musical pulse, structured around relatively short-term and absolute forms of correction.

Pulse in non-musical human interaction

Communicative musicality in parent-infant interaction

Broadly speaking 'pulse' - as a reasonably regular succession of discrete vocal and/or gestural events through time - has been proposed to be inherent in any situation in which two or more individuals coordinate their communications, predictively empathise with their movements, or even just attentively spend time together. Malloch and Trevarthen (2011) argue that it forms, together with 'quality' and 'narrative', a fundamental "communicative musicality" (Malloch 1999). This, they propose, is expressed from infancy onwards and grounds our social understandings, attachments, attunements and the expression of our very vitality itself (Stern, 2004). It is, in their words, "part of a natural drive in human sociocultural learning which begins in infancy" (Trevarthen, 1999, p. 194) - "an expression of our human desire for cultural learning, our innate skill for moving, remembering and planning in sympathy with others that makes our appreciation and production of an endless variety of dramatic temporal narratives possible" (p. 4). In related literature Trevarthen (2016) argues, with reference to his and others' research (e.g. Papousek 1996; Stern 2010; Trehub 1990; Trevarthen 2012; 2015), that "an Intrinsic Motive Pulse of "life time" is the essential property of human action-with-awareness; a foundation for aesthetic appreciation of the intentions and feelings of effort as our bodies try to move well, and also for appreciation of the sympathetic expressions of moral feeling in communication with others" (p. 226). This literature area and the many critical implications for understanding the motivational aspects of communication and music will be discussed further in Chapter 4. Even, however, without delving deeper at this stage we can, I think, already see why one can expect such communicative musicalities to be active and of crucial importance in all human exchanges and particularly so in prelinguistic parent-infant interactions.

Malloch (1999), inspired by finding himself wanting to tap his foot along to recordings, argues for the existence of a pulse in mother-infant communications - the "musical dialogue". He analysed spectrographic representations of recordings and interpreted "bar lines". In one case, for example, he posits that 0.68s intervals emerge as the time unit by which the baby's and mother's utterances may be naturally divided. He also suggests that, in places, groups of three 0.68s intervals form a larger interval of 2.03 seconds which are surrounded by periods of silences. Furthermore, he tries to show occasional correction intervals at 0.34s (a half subdivision of 0.68s) being used to allow parent and infant to "realign". He is suggesting, not just a pulse, but a fully metered structure with coordinative correction mechanisms and summarises his perspective on pulse by stating that "even such an immature human has a complex periodicity of vocal expression comparable with that of adult expression, and is similar to the rhythmic patterns shown by infants in spontaneous vocalisations weeks and months after full-term birth" (Ibid, p. 37/38).

The crucial question with regards to the central goals of this thesis is whether the pulse described is an element of all human communication that allows coordinated companionship to arise and is particularly exploited in music (as proposed by the author). Alternately is there something additional or psychologically/mechanistically distinct in a musical pulse? Although the musical terminology and description is intuitively appealing it seems pretty clear, in my opinion, to be the latter. The

---

[51] We should perhaps note, however, that the description of Gibbons being "monogamous" is much more loosely followed in practice than the term might suggest (see Reichard et al. 2016).

degrees of interpretative freedom in the methodological approach are considerable with bar lines neither reflecting consistently onsets nor perceptual centre. Rather they represent a loosely constructed and retrospective best fit. This should not minimize the findings and psychosocial implications of the analysis nor the full extent of their relevance to understanding aspects of musical pulse and music. However we are qualitatively too far removed, in my opinion, to understand pulse in music as being merely a stricter manifestation of entirely the same process described.

Rhythm in linguistic interaction

Rhythm in speech?

The idea of an isochronous foundation in English speech dates back to the 18th century (see Abercrombie, 1967). Pike (1945) and Abercrombie (1967) presented highly influential theories that languages were either stress-timed or syllable timed and by definition could not be both. Speech was thereby thought to exhibit periodicity at the level of stresses or syllables. Abercrombie added a functional rationale arguing that this was primarily the result of the physiological structure of exhalation. However, a full half century or more of empirical investigation has clearly shown that even those languages viewed as being the most clear-cut manifestations of syllable- or stress-timing do not display isochronous structure on either count (see Bertinetto 1989; Nolan et al. 2015). To mention but a selection of the clearest evidence, studies have variously shown that English stress "feet" lengthen with the number of syllables (Dauer, 1983) and have reported that stress-based and syllable-based languages cannot be discriminated on the basis of the timing of interstress intervals (Roach 1982). A theory of isochronous "morae" (a unit smaller than a syllable) in Japanese (Ladefoged, 1973) has also been debunked (Beckman, 1982; Hoequist, 1983). Experts in the field seem, therefore, to concur that 'periodic' (White, 2014) or 'coordinative' (White et al., 2012) rhythm is absent in speech. Comprehensive and recent reviews conclude unequivocally that we 'have to retreat from any hope that languages are rhythmic in the everyday sense of clock-rhythmic' (Nolan and Jeon, 2015, p. 6) and that progress will only be made "if one thinks of [speech] rhythm as systematic timing, accentuation, and grouping patterns in a language that may have nothing to do with isochrony" (Patel, 2008, p.122). Succinctly, a musical pulse is not generically operational in common speech.

In more recent years the debate has therefore moved to whether or not the flow of speech conforms to more contrastive accounts in which rhythm emerges from the alternation of genrstronger and weaker elements and their relative durations (or even from finer grained distinctions [see Turk & Shattuck-Hufnagel, 2013]). The most influential of the many 'rhythm metrics' or indices that have been proposed, argues that languages can be ascribed a "rhythmic class" based predominantly upon the duration of its vocalic and consonantal intervals. Using a normalised 'PVI' equation focused on the degree of variability between successive acoustic segments or phonological units has been shown to form three clusters of languages which, in turn, appear to correspond to the three traditional rhythm classes (e.g. Ramus et al, 1999; Grabe & Low, 2002; Nespor et al., 2011). This analysis has notably, for our concerns, led to some interesting comparative studies on rhythm perception. Ramus et al. (1999), for example, showed that cotton-top tamarins, like human newborns, have the capacity to discriminate between sentences from languages that are differentiable on the basis of their "rhythm class". Subsequently the role of the rhythmic cues in distinguishing between the classes was supported by studies showing that cotton-top tamarins and human newborns do not differentiate between languages from within rhythmic classes (Ramus, 2002), while 5-month-old human infants do (Tincoff et al., 2005). However, as similar results have been found in experiments with trained rats (Toro, et al., 2003), it seems that the distinguishing characteristics relate to global categories of temporal structure and that the relevant psychological capacities evolved for general auditory purposes early within the mammalian clade. Thus, these studies may be of some interest to the perception of temporally structured events in humans. However, their relevance to rhythm perception is limited and any notion that they indicate an underlying attentional structure that is analogous to a musical pulse is certainly misplaced.

More unpromisingly still there is considerable debate as to whether languages really do fall into distinct rhythmic classes or are really temporally and contrastively structured. The technical particulars have been heavily criticised for being unable to accommodate variations in speech rate, within-speaker differences and for uncertainties in the measuring process (e.g. Gibbon, 2003; Knight, 2011). Nolan and Jeon (2015) summarise that the theory effectively fails to capture the true nature of speech rhythm and they also note "the lack of quantitative evidence from the acoustic signal for a rhythmic dichotomy, and also a lack of empirical support for its reality in the perception of non-linguists and among speakers of different languages (Turk & Shattuck-Hufnagel, 2013; Arvanti, 2009)" (p. 3). Answering the question as to whether we should accept that the relevant notion of rhythm is contrastive rhythm, and whether languages strive to achieve sequential alternation of prominences (albeit with less or more salient 'prominence gradients') they are unequivocal in their interpretation of the evidence stating that: "If so, we must conclude that languages are doing a pretty poor job" (p. 3).

Nolan and Jeon's (2015) view goes further than arguing that there is no teleological goal of rhythmicality in speech. Speech is, they posit, "not so much arrhythmic as antirhythmic – it exhibits a "redolent, wilful and rebellious disregard for decent metrical principles" (p. 7). Timing is, they argue, the servant of linguistic structure, lexical differentiation, and of the prominence for informational purposes and prosodic edge-marking as an aid to parsing utterances (p. 9). As such, rhythm in speech is interpreted to be a constructed metaphorical representation. Instances when speech is aligned to essentially rhythmic structures (such as in music) are thus by extension construed as metaphorical analogies between the rhythm of the target and some non-rhythmically motivated aspect of the structure of the speech. We should, of course, be mindful of the fact that our capacities for creating and perceiving pulse, even a "musical pulse", can be called upon to varying degrees in spoken nursery rhyme, creative and expressive poetry, or in persuasion, or inspirational oratory. However, neither musical pulse, nor a full suite of its constituent capabilities, appears to be a necessary component in everyday linguistic speech.

Turn-taking in conversation

Another possible candidate for behavioural and/or mechanistic analogy with musical pulse is conversational turn-taking. Himberg et al. (2015) appear to be suggesting that turn-taking can represent a full analogy to musical pulse. In this study adult pairs created 5-min stories alternately contributing one word at a time. No instructions as to timing were given and they received either audio or audiovisual connection to each other. Despite the creative challenges of the task and even though the rhythms were 'unstable' (R=0.14 for pooled data), the word rhythms were shown to be entrained (R=0.70, $p<0.001$) to a similar strength as is typically observed in experimentally guided finger-tapping tasks. Thus, the authors argue that "speech seems to spontaneously induce strong entrainment between the conversation partners, likely reflecting automatic alignment of their semantic and syntactic processes" (p.113). This notion accords well with an influential description of an oscillatory model of conversational turn-taking in which syllable production entrainment allows for efficient interlocutor coordination with minimal gap and overlap in talk (Wilson & Wilson, 2005; see also Takahashi et al., 2013). However, the nature of the task in the above study was notably artificial and imposed an unusual degree of regularity. It is debatable therefore whether it reflects more naturalistic turn-taking in everyday linguistic conversation.

Many authors have suggested that the capacities involved in coordinated turn-taking predated language in evolution and provided a necessary pre-condition for linguistic interaction to emerge (see Holler et al., 2015). Considering the literature on 'duetting' in primates (see above) it seems entirely possible that vocal turn-taking is ancestral in origin in the primate order[52]. Levinson (2016) suggests that it can be viewed to "belong to a package of underlying propensities in human communication, including the face to face character that affords the use of gesture and gaze, and the motivation and interest in other minds – 'the interaction engine' (Levinson, 2006; Tomasello, 2009)" (p.10). In general, turns in conversation do indeed exhibit certain foundational charactistics (Sacks et al., 1974; Levinson & Torreira, 2015) and a remarkable consistency across cultures (Stivers et al., 2009). Turns tend to be brief (roughly 2 seconds on average) with occasional efforts for a longer turn of narration. Overlap between individuals (defined as an overlap less than 100ms) occurs very infrequently and typically accounts for less than 5% of the speech stream. The system, furthermore, easily accommodates a wide range in the numbers of speakers and seems to be a centrally operational and modality non-specific structure. The modal gap of roughly 200ms between turns, for example, remains constant in the absence of visual information (Levinson & Torreira, 2015). The general structures are, after accounting for necessary preparatory and held movements, also exhibited correspondingly in sign language interaction (de Vos et al., 2015). It is also important to note that turn-taking proto-conversations with caretakers in pre-linguistic early infancy (e.g. Bruner, 1975) show similar timings and structure except for exhibiting, on average, greater overlap (see Gratier, 2015).

Turn-taking is notably a complex social, psychological and cognitive task, operating at the limits of human performance (Levinson, 2016). The cross-linguistically consistent modal response time of 200ms (Stivers et al., 2009; Levinson & Torreira, 2015; Heldner & Edlund, 2010) would represent a decent, if unremarkable, reaction time for a primed single action. However, it is much more impressive when we consider the range of possibilities in conversation. Preparation across languages begins 600 or 1000 ms before speech of a single primed or non-primed word respectively (Indefrey & Levelt, 2004; Indefrey, 2011; Bates et al., 2003), and around 1500 ms before uttering a short clause (Griffin & Bock, 2000) (see Roberts et al., 2015). Therefore, given that turns on average last only 2 seconds, responses need to be prepared already halfway between the start and conclusion of the conversational partner's previous turn (Levinson & Torreira, 2015).

Consequently we can safely assume that a considerable degree of temporal prediction is involved. However, although it is tempting to imagine that these predictions are based on temporalities - on a pulse - it seems that other indications predominate. First of all an understanding as to the function (i.e. is it a question or statement?) of the others' speech act occurs within 400 ms of the beginning of the turn (Gisladottir et al., 2015; see Levinson, 2013). It starts to trigger initial production processes within a further 500ms (Bögels et al., 2015; but see also Sjerps and Meyer, 2015). Unsurprisingly the person planning the onset of his or her turn can also garner predictive information from the lexical, semantic, and syntactic structure of the partner's utterance (Magyari et al., 2014, Garrod and Pickering, 2015). He/she may even prime some educated guesses of upcoming words (Magyari and de Ruiter 2012) based largely on the semantic context (Riest et al., 2015). Finally, some more intrinsically temporal and prosodic cues are also available from, for example, lengthened syllables towards the end of a turn (Bögels & Torreira, 2015). These processes and the many fascinating coordinative processes involved in turn-taking and conversation are discussed and reviewed more extensively elsewhere (e.g. Holler et al., 2015; Levinson, 2016). The crucial point for our purposes, however, is that the predictive mechanisms are comparatively short-term with no direct evidence for a mutually understood and continuing "musical" pulse across turns.

Musical Pulse and Rhythm in Development

In this chapter I have deliberately focused attention on the most basic manifestation of a musical pulse and its corresponding mechanisms, capabilities and embedded connections. This has been partly to constrain the boundaries of investigation but also to keep the discussion, as much as possible, at the level of a shared human capacity for music and away from culturally constructed learning and convention. Meter is an interesting case, however, in that the principle of operating within a meter undoubtedly relies on established mechanisms of a musical pulse. It is also a music-specific trait that shares some common ground across cultures in being interpretable at different hierarchical levels and gestalts (see London, 2012 for a full discussion). That said it is also manifest in a considerable degree of cultural learning and appreciation. Very broadly

---

[52] It is a notable however that great apes seem to prioritize gestural communication systems (Call & Tomasello, 2007) and have not been documented in exhibiting vocal turn-taking

speaking, we seem to come into the world primed to adopt a complete variety of metric structures and then gradually get constrained by cultural experience. Infants are already sensitive to pulse and meter in musical stimuli and will, after habituation to three similarly metered patterns remain attentive to a novel rhythm that embodies a different meter for longer than to one with a similar metric structure (Hannon & Johnson, 2005). Their relative flexibility compared to adults was clearly showed in a well-designed study by Hannon and Trehub (2005a). Adults from North America, Bulgaria and Macedonia were tested on their ability to distinguish between “simple” or “complex” metered Bulgarian and Macedonian folk tunes and meter-preserving and meter-disrupting variants. North American adults, reflecting the dominance of simple meters in Western music, were unable to distinguish between meter-disrupting and meter-preserving variations on complex metered tunes. Bulgarians and Macedonians, who are acculurated to both forms of meters, were successful at both tasks. The notion that North American adults gradually lose their ability to respond to complex meters rather than Bulgarian and Macedonian acquiring new skills was evident in the finding that 6-month-old infants prefer variations that disrupt the original meter to those that preserve it for both simple and complex meter tunes. Culture-specific biases would seem to appear between months six and twelve in ontogeny as North American 12-month-old infants start to perform worse on complex-meter tunes (Hannon & Trehub, 2005b). Although at this age, in contrast to adults, this can quickly be readdressed with exposure to Bulgarian folk tunes for 20 min daily for two weeks (Hannon & Trehub, 2005b).

Similar points could perhaps be made regarding correlations between rhythmic structure and broader gestalt principles and auditory scene analysis capabilities from early development. Musical rhythms can, of course, be created in an almost infinite variety and according to a whole world full of compositional traditions. Some of the principles of construction might, however, be argued to reflect basic grouping capacities exhibited from early infancy. 4- to 6-month-old infants have been shown to perceptually group or phrase events in studies showing that they more accurately detect pauses after short notes than long notes (Trainor & Adams, 2000). They also listen to Mozart minuets for longer if pauses are injected at phrase boundaries (and are marked by note lengthening and falling pitch) than elsewhere (Krumhansl & Jusczyk, 1990; Jusczyk & Krumhansl, 1993). Thus they would seem to have expectations as to which structural features are likely to signal phrase boundaries. It is also worth noting here (although not a purely rhythmic phenomenon) that an early form gestalt formation or instance of auditory scene analysis (Bregman, 1990; 1993) in music is shown in a study reporting that 6- and 8-month-old infants find it easier to detect timing alterations within tone groups than between tone groups (Thorpe & Trehub, 1989).

Chapter Conclusion – A Musical Pulse

A musical pulse is in practice, of course, more fascinating, variable, playful, and complex than the scope of the discussion and studies considered above. "Finding" or expressing the beat in music involves an intricate understanding of musical structure – in some sense an intrinsically learned and manifest statistical analysis – and cultural knowledge and acculturation. Stobart and Cross (2000), for example, reported an interesting study of the Easter Songs of Northern Potosí, Bolivia in which the appropriate understanding of whether the pulse was "on" or "off" the beat was entirely culturally contingent. These kinds of issues are undoubtedly integral to the experience and function of music and to our understanding of social intentions, individualities, intrinsic learning and cultural forms. However, here the focus has deliberately been on the "bare bones" – those "basic" processes and capabilities that allow us to create and interact with any musical pulse and thus root any particular manifestations and cultural quirks. We can thus imagine that these reflect universal root capacities that underpin any engagement with pulsed music across time and cultural divides. They are effectively the architectural groundwork upon which individual and cultural variabilities can be constructed.

Pulse in interaction has, in some writings been described as a continuum with isochrony defining one end of the scale and with music positioned towards that extreme (Cross et al., 2011; Koelsch, 2012; cf. Bispham 2006). According to this view there is in this respect no definite distinction between music and language with areas of overlap, and transitional zones from isochronous to non-isochronous signals in both music and speech. The "illusion" in which repeated speech patterns will be interpreted as being "musical" over time has sometimes been seen to support this nebulous divide between "pulse" in language and music (Deutsch et al., 2011). However, this could also be interpreted as a parsimonious move from one mode of processing to another based upon the amount of inherent temporal regularity in the signal. Describing the shift towards a "musical" pulse in terms of isochrony and therefore in terms of a relatively absolute physical characteristic can, I think, be misleading. It reflects neither the reality nor the *primary* goal of a musical pulse. As we have seen, a musical pulse can, by virtue of a wide range of involuntary fluctuations and social and learned cues, be both temporally flexible and "stable". Thus, I would suggest, a more primal way to consider the specificities of a musical pulse is to consider the involvement of second-order relational processing over time, its inherent continuity, and correction mechanisms. It is these features that profoundly create the stability of a musical pulse. Relatively high degrees of behavioural isochrony are a particularly likely resultant attribute. I propose therefore and conclude from the comparative discussions throughout this chapter that a musical pulse is distinct in the integral and particular nature of sustained attention; in its continually operational period correction mechanisms; a corresponding degree of volitional control; and an awareness of the underlying temporal beat.

In essence a musical pulse is a particular type of attentional and temporal framework that allows the management of individual experience (and of smooth and ballistic fine motor timing) and a mutually manifest coordinating structure for interpersonal interaction. It uniquely affords, and is most strikingly and evidently displayed in the harmonised coordination of action and perception in ensemble performances. These are often described as impressive accounts of group synchronisation. However, to be specific and focus discussion on the root enabling features of musical behaviour, we should perhaps be clear in describing action coordinated with reference to a shared and persistent attentional pulse. The psychosocial, intersubjective, motivational and functional implications inherent in the specificities of a musical pulse are obviously intense both in terms of the flow of individual phenomenology and the social and communicative effects of sharing musical time with others. These will be discussed in chapter 4.

A musical pulse is a sustained, volitionally controlled, future-directed attentional pulse, within a temporal window of the psychological present. It incurs awareness of a pulsed framework and is perceived unambiguously or at hierarchically related levels by acculturated individuals. Furthermore, it involves engagement of the motor system in a way that enables management of fine and gross temporal control in ballistic and smooth movements, and mechanisms for period correction based upon the output of self and/or others.

Chapter 3 – Musical Tone

This chapter mirrors the approach taken in chapter 2 in aiming to identify features that are specific to a sustained musical tone. The use of pitch and tone in communication is widespread in human and non-human interaction. I will therefore be aiming for a more intricate distinction in describing unique psychophysiological features of musical pitch perception and production. In the opening section I offer a comparative discussion of pitch perception. This highlights the increasing amount of evidence in support of the view that the most fundamental features of pitch perception are neither unique to music or indeed to our species. They are, broadly speaking, widely shared with other primate species.

The Psychophysiology of Pitch Perception

Pitch, like pulse, is ultimately a perceptual construct. It is relatable to physical frequency (assuming the F0 is within a range of about 30 Hz to 5000 Hz) but is by no means a direct perceptual correlate thereof. It is a component in our broader suite of auditory scene analyses (Bregman, 1990) and, as such, one way in which we group related partials and make sense of the complex auditory stimuli we encounter in music, language, and in general interaction with the external world (See Darwin, 2005 [in Plack et al. 2005]). The functioning and physiology of the auditory system as well as models of frequency analysis and tonotopic coding through the basilar membrane in pitch perception are well-developed and fascinating research areas. Recent and extensive reviews are available in the literature (e.g. Cheveigne, 2010; Oxenham 2010; 2012; McDermott & Oxenham, 2008; Plack et al., 2014; Stainsby & Cross, 2016), so I will focus here on the most critical behavioural features of pitch perception in humans and respective comparative studies in other species, focusing particularly on primates.

Philosophically speaking we do not need to assume *a priori* that pitch is a pertinent percept in other species. However it does seem that it is indeed widely operational across taxa. Studies of frequency discrimination liminality (FDL) in humans and primates, using mostly pure tone stimuli, have shown successful discrimination in chimpanzees (Kojima, 1990), Old World monkeys (Stebbins, 1973; Sinnott, 1985; 1987; 1992; Prosen et al., 1990) [~1.2% change in frequency], and New World monkeys (Capps and Ades, 1968; Recanzone et al., 1991; Wienicke et al., 2001; Osmanski et al. 2017) [~2.5%]. FDL is notably lowest in humans (Wier et al., 1977; Sinnott et al., 1985; 1987; 1992) [(~0.2%]. However, all species tested showed lower FDLs with frequencies associated with their best hearing sensitivity (e.g., Stebbins,1973; Wier et al., 1977; Kojima, 1990; Wienicke et al., 2001) and a correlation between low FDL and a concentration of spectral energy in that species' vocalizations (see Winter et al., 1966; Newman, 2003; Osmanski et al., 2017). This clearly suggests that pitch is communicatively and socially relevant in primates and that auditory perception and vocal production are evolutionarily interrelated (Kuhl, 1988)[53].

Most, but by no means all, sounds that evoke a pitch in natural settings have a harmonic spectral structure in which acoustic salience is concentrated at frequencies that are integer multiples (harmonics) of a common fundamental frequency (F0). The purest evidence that pitch represents a perceptual construct resulting from analysis of complex physical stimuli (see McDermott & Oxenham, 2008) comes from well-established evidence that the percept of pitch remains stable following the removal of the actual physical F0. EEG studies have shown that, in humans, the ability to find the pitch of a missing fundamental emerges between 3-4months (He & Trainor, 2011). It is behaviourally well established by 7 months (Clarkson and Clifton, 1985; Clarkson and Rogers, 1995; Montgomery and Clarkson, 1997). This capacity appears to be common across species and has been experimentally documented in birds (Cynx and Shapiro, 1986), cats (Heffner and Whitfield, 1976), rhesus monkeys (Tomlinson & Schwarz, 1998), and chinchillas (Schofner, 2011). Furthermore, evidence for a shared neurological process in monkeys and humans comes from reports of a shared specialized area for pitch processing in the low-frequency region of the auditory cortex (Patterson et al., 2002; Penagos et al., 2004; Bendor and Wang, 2005, 2010; Schneider et al., 2005; Schönwiesner and Zatorre, 2008; Puschmann et al., 2010). Pitch-selective neurons in this region in marmosets (*Callithrix jacchus*) can encode the pitch of missing fundamental (MF) sounds that are spectrally outside their excitatory frequency response area (Bendor and Wang, 2005; 2006) and are notably sensitive to pitch salience and temporal envelope regularity (Bendor and Wang, 2010).

In humans, a sense of pitch may also arise from stimuli with inharmonic[54] components (Shouten et al., 1962), spectra comprised of broadband noise (Burns and Viemeister, 1976; Cramer and Huggins, 1958) and/or from a psychological composite of stimuli to different ears (Cramer and Huggins, 1958; Houtsma and Goldstein, 1972; Klein and Hartmann, 1981). More recent studies of pitch-matching and melody discrimination have also revealed that the pitch of the missing

---

[53] Genetic and anatomical comparative perspectives confirm differences in the anatomical structures of the outer and middle ear between chimpanzees and modern-day humans consistent with co-specific vocalisations (Clark et al., 2003; Abdehak et al., 1997; Vervoort et al., Masali et al., 2006; Nummela, 1995). Empirical data of hearing capabilities show distinctions between the two species. Chimpanzees are optimally sensitive at 1 kHz (corresponding to the most salient information of species-specific pant-hoots) and 8kHz (Kojima, 1990). Human sensitivities differ in that, while chimpanzees exhibit a loss of sensitivity between 2 and 4 kHz, we maintain a high sensitivity in this region that contains relevant information in human vocalization (Jackson et al., 1999; Brown & Waser, 1984). Inspired by this discrepancy Martinez et al (2004) analysed, using a physical model, the influence of skeletal structures on the acoustic filtering of the outer and middle ears in five fossil human specimens for the middle Pleistocene site of the Sima de los Huesos (SH) in the Sierra de Atapuerca, Spain. Their analysis concluded that the skeletal anatomy of SH hominids of the outer and middle ear is "compatible with a human-like sound power transmission pattern" and clearly different from chimpanzees in the critical region of ≈ 4kHz. Furthermore they infer that, as SH hominids form part of the Neanderthal evolutionary lineage (Arsuaga et al, 1997; Martinez & Arsuaga, 1998), this condition may well have been present in the last common ancestor of modern humans and Neanderthals – *Homo antecessor* (Bermudez et al., 1997).

[54] Inharmonic refers to instances in which the harmonic partials (frequencies) of a sound source are not related by small integer values (most likely producing a 'rough' sound)

fundamental can be extracted even when all the harmonics present are well above 5 kHz (Oxenham et al., 2011). These and many other findings have largely been interpreted within general pattern matching or temporal models of pitch perception (see Cheveigne, 2005; 2010). They "principally differ in the relative importance of frequency components that are resolved (RES) or unresolved (URS) by the auditory system" (Stainsby & Cross, 2016, p. 68 {abbreviations added}; [see also Plack and Oxenham, 2005; Plack et al., 2014]).[55] Essentially pitch can be evoked in three ways: By the use of the lowest-frequency component present in a harmonic sound or a pure tone (Ohm, 1843); by matching spectral harmonic templates in the central auditory system to the resolved components of an incoming harmonically structured sound (Goldstein, 1973; Shamma & Klein, 2000); or by extracting pitch from temporal interactions between unresolved frequency components within an auditory filter - the 'temporal envelope' (Schouten, 1940). Although each process reflects mechanisms that can operate independently, the relative strength of the perceived pitch - measured inversely as differences in the ability to discriminate changes in F0 (Micheyl et al., 2010) - varies[56].

In particular, studies of F0 difference limen (FODL) have shown that complex tones have greater pitch strength than pure tones (Zeitlin, 1964; Henning & Grosberg, 1968; Fastl & Weinberger, 1981; Spiegel & Watson, 1984) and even that the pitch strength of harmonic complex tones is unaffected by removal of the fundamental frequency (Schouten, 1938; Licklider, 1956). Additionally, we see a reduction in pitch strength as spectral content moves beyond the lower range of resolved frequencies. Thus, in general, RES seem to dominate over URS. We also seem to demonstrate a stronger pitch percept to sounds in which the RES are spectrally harmonic (Faulkner, 1985; Moore & Glasberg, 1990; Micheyl et al., 2012) and sounds in which the salience of temporal envelope cues in URS is high (Houtsma & Smurzynski, 1990). As mentioned above much more detailed accounts are available in the literature. However, from a comparative perspective it has been crucial to outline these central findings in humans as Shofner and Chaney (2013) have proposed that pitch perception in non-human animals is much more restricted. They propose, with reference to a study on Chincillas that pitch perception results exclusively from temporal, rather than spectral, processing in nonhuman mammals. This suggestion has, however, been strongly countered in a pertinent study by Song and colleagues (2015). After considering differences in resolvability boundaries between humans and marmosets (presumably due largely to cochlear size variance [see Osmanski et al., 2013]) they found that this highly vocal species shares with humans all three primary mechanisms/features of central pitch processing discussed above. They conclude, therefore, that marmosets and humans share similar pitch perception mechanisms and suggest further that these mechanisms may have emerged early in primate evolution. As marmosets are a relatively distant primate cousin (with a last common ancestor circa 40 million years ago) this does seem quite plausible. However, we do perhaps need more similar studies on other primates to confirm this hypothesis over, for example, a homoplasic evolutionary scenario[57]. Nevertheless, from the results of the studies mentioned above, it does seem that we can already be confident in accepting Song and colleagues' conclusion (2015) that our basic psychophysiological mechanisms of pitch perception are, at least broadly speaking, shared with other primate species.

Relative Pitch Processing

This section explores the psychology of relative pitch processing. In particular I will show that it emerges early in human development and is widely operational across communicative domains in humans. However, similarly to relative/beat-based timing (chapter 2), it represents a stark contrast between human and non-human cognition and perception.

As noted by McDermott and Oxenham (2008, p.453) "standard views of the auditory system might lead one to believe that absolute pitch would dominate perception". They note, for example, that "tonotopic representations that are observed from the cochlea [Ruggero, 1992] to the auditory cortex [Kaas et al., 1999] make absolute, rather than relative, features of a sound's spectrum explicit". However, in humans, relational pitch processing is ubiquitous in perception and communication and is effortlessly acquired in early development. In music, while absolute cues are certainly operational (see Loui, 2016), it is a relational architecture that principally and primarily characterizes global (e.g. tranposability) and local (e.g. intervals) pitch systems across cultures[58]. It certainly, therefore, does not make sense to argue, as Mithen (2005) does, that the acquisition or retention of absolute pitch makes man or Neanderthal "more musical" (see Bispham, 2006a). If anything precisely the opposite is the case.

Despite some debate in the literature, the majority of the evidence appears to support the notion that infants prefer to process pitch combinations in terms of the relationships between tones rather than absolute frequency values (Trainor, 2005). Saffron and colleagues have, however, argued to the contrary with a series of statistical learning studies in which infants

---

[55] Resolved harmonics at those (~5-10) lowest frequencies that are separated into different auditory filters and can be heard individually from the whole complex sound (Plomp & Mimpen, 1968). These harmonics are known as resolved harmonics (RESs). Unresolved harmonics conversely are those for which each auditory filter receives significant power from more than one filter. The boundaries of RESs and URSs have been determined in behavioural studies (e.g. Shackleton & Carlyon, 1994; Bernstein & Oxenham, 2006).

[56] It is interesting to note that in congenital amusia, individuals have been shown to have a specific alteration in the processing of resolved harmonics (Cousineau et al., 2015). Fine-grained pitch perception difficulties have also been shown to extend to 8 kHz which is notably beyond the typical region of musical pitch (Whiteford & Oxenham, 2017). This study also showed some surprising deficits on amplitude as well as frequency distinctions.

[57] Another area that remains to be investigated comparatively, as pointed out in Patel & Demorest (2013) is the right-hemisphere bias in missing fundamental processing in humans (Patel & Balaban, 2001; Zatorre 1988). Following Zatorre et al. (2002) they suggest that this may reflect a trade-off in specialization between the right and left auditory cortices, with enhanced spectral resolution and temporal resolution in the right and left hemispheres, respectively (see Pöppel, 2003).

[58] Will (1997) notes an intriguing partial exception in Central Australian songlines in which 'intervals in the upper melodic range are shown to be *linear* shifts of intervals in the lower range' (emphasis in the original, p.6)

receive continuous exposure to sequences of tones and are subsequently tested on their recognition of the most frequent three-tone fragments. The key conclusions taken from these studies were that, in the presence of both relative and absolute cues, infants prefer absolute cues both in randomly selected twelve-tone and standard seven-note major scale samples (Saffron & Greigentrog, 2001; Saffron, 2003). In the absence of absolute cues they are, however, able to learn on the basis of relative cues (Saffron et al., 2005). Further evidence that infants can relate to features of absolute pitch comes from studies showing that seven month old infants, similarly to college students remembering soap signature tunes (Schellenberg & Trehub, 2003), retain absolute pitch information about unfamiliar lullabies after two weeks of exposure (Volkova, 2004 [reference taken from Trehub & Hannon, 2006]). We should perhaps note that the results of Saffron's studies may be influenced by the amusicality – structurally and timbrally - of the stimuli presented. In terms of memory for real tunes, six month old infants recall, after relatively minimal exposure, relative pitch features (Trainor, Wu, & Tsang, 2004) which indeed appear to be primarily salient (Platinga & Trainor, 2005). In shorter-term experimental paradigms with musical stimuli, five to ten month old infants have been shown to treat transpositions of original melodies as familiar or equivalent even after limited exposure. Similarly to results in adult studies (e.g. Trehub et al, 1987), infants respond most readily to changes in contour (Chang & Trehub, 1977; Trehub, Bull, & Thorpe, 1984; Trehub, Thorpe, & Morrongiello, 1987) and often confuse altered melodies that maintain shared pitch contours (e.g. Trehub, et al, 1987; Trehub, Thorpe & Trainor, 1990). Markedly the salience of pitch contour is not restricted to musical stimuli but also appears in engagement with infant-directed speech (e.g. Fernald, 1991). This is perhaps not surprising given that, as argued by McDermott & Oxenham (2012), “the difficulty in encoding intervals might simply be due to their specification requiring more information than that of a contour” (p. 3). Significantly infants do seem, however, to be able to differentiate same-contour melodies consistently if the original melodies are based primarily upon low-integer intervals or conform to certain “grammatical” rules (Cohen et al., 1987; Trainor & Trehub, 1993b).

In stark contrast nonhuman animals have so far been shown to have, at best, very limited relative pitch capacity. Songbirds rely instinctively and extensively on absolute cues in recognising sequences of tones (Hulse, 1992; Weismann et al., 2006). Even after vast training they make very limited use of relative cues (Hulse & Page, 1988; Page et al, 1989; Weisman et al., 2004). Although songbirds struggle to recognize transposed isochronous tone sequences, Bregman and colleagues (2012) demonstrated that European starlings (*Sturnus vulgaris*) can be trained to recognize frequency-shifted versions of songs from other conspecifics. It seems likely, however, that the birds may recognize the transposed versions of the songs on the basis of timbral and rhythmic relations (e.g. Hulse & Klein, 1993). A subsequent study (Bregman et al., 2015) demonstrated that, in this species, small manipulations altering either pitch or timbre independently result in an inability to recognize melody. However, melody differentiation generalized even in the absence of pitch if the spectral shapes of the constituent tones were maintained. It therefore seems likely, as proposed by Patel and Demorest (2013), that “the difficulty birds have recognizing transposed tone sequences reflects a general difficulty that animals have with recognizing sound sequences on the basis of relations between acoustic features (McDermott, 2009)” (p. 655). This can be related to the lack of capacity in many other mammalian species to manipulate pitch, timbre and loudness in vocalization individually, and a corresponding lack of relevance in isolating these features in the assessment of conspecifics communicative signals (e.g. Bispham, 2009; Patel et al., 2017).

One of the most significant features of relative pitch in music is that of octave equivalence – the sense that pitches separated by an octave (a 1:2 frequency relationship) are in some way the same. In many cultures it comes naturally to sing in octave unison and the percept of octave equivalence has been demonstrated in speech (Peter et al. 2008), music (Burns, 1999; Thompson, 2013) and even in human infancy (Demany & Armand 1984). That said some cultures sing more spontaneously in internals other than the octave[59] and some studies deny equivalency in 4-9 year old children (Sergeant 1983) and non-musicians (Allen 1967). Thus we should, perhaps, not be too unequivocal in supposing music theoretic accounts of direct equivalency. Unsurprisingly, no strong octave relevance is found in European starlings, (Cynx 1993; Hoeschele et al., 2012). A study on chickadees similarly suggests that, despite some seemingly positive findings, they use pitch height perception to classify pitches into ranges (independent of the octave) to transfer note-range discriminations from one octave to the next (Hoeschele, 2013). In mammals, octave equivalence has rather unconvincingly been reported in rats (Blackwell and Schlosberg 1943) and, in a more significant set of studies, in rhesus macaques (Wright et al., 2000). The latter reportedly retained an identity of tonal melodies (but not synthetic ‘melodies’) if, and only if, they were transposed in whole octaves. Notably large stretches of training were required and we should perhaps also be cautious in interpreting the findings as some spectral absolute cues may have been retained in the harmonics of the keyboard voices used (i.e. a fundamental frequency may have been a first harmonic in transposition). In a study of non-octave single pitch intervals in ferrets, substantial training was also needed to show some recognition of similarity following transposition (Yin et al., 2010). Again this shows that relative pitch in this species is, at best, not a natural process and highly limited.

Relative pitch processing is crucial in all forms of human communication. We can perhaps posit that music is distinct in organizing relative cues in discrete pitches, harmonic blending, and more fine-grained divisions (see below). Nevertheless, relative characteristics are more broadly perceptually and functionally relevant from early infancy onwards. For example, infant-directed speech ((Fernald et al., 1984; Papousek et al., 1985) and lullabies (Unyk, et al., 1992; Trehub et al., 1993; Trehub et al., 1998) consistently employ descending contours in soothing and calming (see Bispham, 2002). In mature speech too, we need to express and interpret functionally significant intonational patterns or even common pitch contours and relations between pitches in tone languages (see Ladd, 2008). More generally still, in relating and physically empathizing with others, we need to account for interpersonal differences in size, physiological structure and subsequent

---

[59] Blacking (1995), for example, states that “Venda people would resist singing in unison with [him] when asked to do so, on the grounds that with more than one person present any good musician would prefer to sing in harmony at the interval of a fourth, fifth, or third” (p.236).

differences in vocal range and tessitura. Hence pitch can perhaps be seen as another communicative cue in which relative processing evolved to enhance social understanding and communality. Notably, as a group, autistic individuals attend more to absolute pitch cues than normal individuals. This has been shown both in speech and music and it has been suggested that this is one basis of their communicative challenges in language (Heaton, 2009; Heaton et al., 2008; Järvinen-Pasley et al., 2008; Järvinen-Pasley et al., 2008). Relative pitch, similarly to relative/beat-based timing, appears to be a major transition between humans and other mammalian and primate species. It is an essential component in the capacity to engage with a musical tone without being intrinsically specific to it. It is, more broadly, a critical and generalisable feature of human communication. In evolutionary terms, it seems entirely plausible, therefore, that broad relative pitch capacities initially evolved and developed independently of music.

Perceptual Consonance

This section continues the comparative approach exploring whether the percept of consonance and dissonance is shared with non-human species and whether it is mechanistically distinct from perceiving degrees of timbral roughness in natural sounds. In particular, I will argue against the notion that music can be cross-culturally characterized by an aesthetic preference for consonant intervals. Processing dissonant intervals does seem, however, to incur a greater attentional cost in humans (a feature not found in the relatively few available studies on non-human species).

A connected perceptual concept to relative pitch processing is that of consonance and dissonance. Typically intervals, presented either as melodic sequence or horizontal harmony, are described and perceived as being progressively more consonant as fundamental frequencies form more simpler (i.e., small-integer) ratios. Categorical distinctions between levels of dissonances appear early in human development (Schellenberg & Trainor, 1996; Perani et al., 2010) and distinct neural responses have been observed in the inferior colliculus and the primary auditory cortex of cats and monkeys respectively (McKinney, et al., 2001; Fishman et al., 2001). Behaviourally, both pigeons (*Columbia livia*) and Japanese macaques *(Macaca fuscata)* have been successfully trained to discriminate among chords based on sensory consonance (Brooks & Cook, 2010; Izumi, 2000). Consonant intervals are often described as being more pleasant to the ear, and a desire for consonance has variously been argued, implicitly and explicitly, to underlie the structure of musical form and tonality. Indeed, if we restrict our view to Western tonal music written between c. 1650-1920 this perspective can appear to find a certain degree of music-theoretic, and even neuroimaging support. Bidelman and Grall (2014) have, for example, argued that listeners' behavioural preferences for intervallic and chordal harmonies are well predicted based on underlying sensory pitch representations along the early auditory pathway (rostral brainstem and auditory nerve) (Bidelman and Heinz, 2011; Bidelman and Krishnan, 2009). Even when acoustic or peripheral factors are controlled, brain activity provides, they posit, a sufficient basis to account for the fundamental organizational principles of tonal music (Bidelman, 2013; Bidelman and Krishnan, 2009; Itoh et al., 2010). They conclude that "these studies demonstrate that nascent features of tonal structure and musical pitch hierarchy are present in the earliest (pre-attentive) stages of the auditory system in neurocomputations operating below conscious awareness" (p. 205). This may be partially reflective of a manifestation of functional concepts of object and beauty in Western 'art' music (see Bohlman, 2000). Even here, however, the idea is musicologically speaking highly problematic, not least because it is, more often than not, the interplay *between* relative degrees of dissonance that is more significant than any absolute degree of consonance (Meyer, 1956; Krumhansl, 2000; Woolhouse & Cross, 2010). As we attempt to widen the scope of relevance to ethnomusicological and cross-cultural concerns, the idea falls apart completely (e.g. Vassilakis, 2005; Jordania, 2006). This is important to clarify as misrepresentations of aesthetics, biology and consonance in music abound in the literature. At the centre of the difficulties encountered is a need to avoid conflations of issues regarding spectral harmonicity, intervallic consonance, processual predispositions, biological parsimony, musical structure, and cultural conventions and aesthetics. Distinctions (and partial correlations) between the above are theoretically well-known but nevertheless arguments, for example, regarding musical aesthetics and a biological basis for musicality still result from talking at cross purposes (e.g. McDermott et al., 2016; Bowling et al., 2017).

Varying degrees of periodicity and roughness are inherent in the timbre of natural sounds and the organization of spectral harmonics. In general communicative signals and vocalisations are most likely to exhibit high levels of harmonicity. So it is perhaps not surprising that primates, being social animals, exhibit significant perceptual sensitivity to the harmonic relationships that define consonance as well as distinct neuronal responses (Feng & Wang, 2017; Fishman et al., 2013; Wang, 2013). Some prominent theoreticians have intuited and proposed that the physical nature of harmonics, vocalization and sensitivities to intra-species communicative intent, may root the psychoacoustic relevance of consonance and dissonance in music (e.g. Rameau [see Christensen, 2004], Terhardt, 1984). In practice, levels of spectral 'harmonicity' and tonal 'consonance' will naturally co-occur as the harmonics of two or more natural musical tones create capricious gradations of auditory 'roughness' or 'beating' in the cochlear mechanism (Plomp & Levelt, 1965; Bidelman & Heinz, 2011). However, in experimental settings they can be artificially separated predominantly by using sine tones and/or presenting tone combinations dichotomously (Houtsma and Goldstein, 1972; Bidelman and Krishnan, 2009). Using controlled stimuli and methods of presentation, a connection does, nevertheless, remain partially inherent in evidence that harmonicity is correlated with preferences for consonant over dissonant chords (McDermott et al., 2010; Krumhansl and Cuddy, 2010). Furthermore, Cousineau and colleagues (2012) showed that impairment in consonance perception commonly found in participants with amusia is correlated with a diminished ability to perceive harmonicity.

It seems now quite clear that Helmholtz's (1885/2013) conclusion that auditory roughness is the "true and sufficient cause of consonance and dissonance in music" (p. 227) is not supported by subsequent empirical evidence. The studies cited above also showed that the strength of consonance preferences did not correlate to the intensity of aversion to roughness in normal subjects (McDermott et al., 2010). Furthermore, despite not describing customary aesthetic considerations for consonance and dissonance, amusics performed normally in terms of the perception of roughness (Cousineau et al., 2012). Neurological studies also exhibit distinct correlations between tonal dissonance and roughness. For example, lesion studies have indicated that the perception of dissonance and roughness can be selectively impaired independently of the other (Tramo et al., 2001). In this light, it is interesting to note also that human auditory cortical brain regions involved in pitch perception and those involved in the analysis of spectral shape have been shown to be partially separate (Norman-Haignere, et al. 2013; Warren, et al., 2005).

The main source of confusion in comparative and cross-cultural accounts of consonance and dissonance comes from a conflation of processing predisposition and preference. As humans we generally find consonant/low integer interval easier to produce and process (Trainor, 1997; Trainor et al. 2002; see Trainor & Corrigall, 2010). Adults and 9 month old infants, for example, detect changes in repeating intervallic patterns or melodic patterns more easily if the stimulus is composed of consonant intervals compared to dissonant intervals (Schellenberg & Trehub, 1994). Sensitivities to changes in tuning are more sensitive to consonant sequences from early on in development (Schellenberg & Trehub, 1996). The processing advantages have also been shown to extend beyond the constraints of a particular task thus suggesting a greater attentional load in perceiving dissonant material. Masataka and Perlovsky (2013) have, for example, shown a greater interference effect of dissonant stimuli and Komeilipoor and colleagues (2015) found that performance in a sensorimotor synchronization task was better after the presentation of consonant stimuli than after the presentation of dissonant stimuli. Notably rats, who have been successfully trained to discriminate consonant from dissonant sequences but not to generalize such discrimination to novel chords (Crespo-Bojorque & Toro, 2015), show no evidence of a processing advantage for consonant structures (Crespo-Bojorque & Toro, 2016; see Toro & Crespo-Bojouque, 2017). It would certainly be interesting to investigate whether primate or other more vocal/social species shared a greater ease in processing consonant stimuli. If so it might suggest an evolutionary connection to spectral harmonicities found in vocal communication. In terms of humans and music, these processing advantages have been argued to parsimoniously account for a statistically high use of low-integer intervals and, for example, certain widespread commonalities in scale structures across cultures (see Burns, 1999; Thompson, 2013; Brown & Jordania, 2013; Bowling, 2015; Bowling et al., 2017). However, extending the argument to hypothesise a general pattern of aesthetic preference for consonance in music across styles and cultures does not inherently make sense and fits neither observational, music theoretic, nor experimental evidence.

It is true that some studies have shown descriptions of preference for single consonant chords over single dissonant chords across cultures (Butler & Daston, 1968). Furthermore it has often been argued that infants, even two-day-old infants born to deaf parents (Masataka, 2006), prefer consonance to dissonance (Trainor & Heinmiller, 1998; Zentner & Kagan, 1998). This contrasts with studies on cotton-top tamarins (*Saguinus oedipus*) and campbells monkeys (*Cercopithecus campbelli*) who displayed no locational preference for being near consonant versus dissonant chords or sounds (McDermott and Hauser, 2004; Koda et al, 2013)[60]. Notably, however, 6 year old human children (unlike 9 year olds) do not describe an aesthetic preference (Valentine, 1962). Furthermore, Plantinga and Trainor (2014) have questioned whether infant looking times in the above mentioned studies necessarily indicate preference as opposed to some other form of categorical distinction. In their own study with 6-month-old infants they showed that, after short exposure, infants did not show distinct behaviours to consonant or dissonant stimuli. Instead the infants simply paid more attention to the stimuli with which they were more familiar. An effect of familiarity, and perhaps also of a notion that dissonance is partially perceived in relation to levels of processing complexity/challenge, is also evident in adults whose pitch processing is facilitated by familiarity. This, in turn, also affects positively the perception of a chord being consonant (McLachlan et al., 2013). In a recent cross-cultural study the Amazonian Tsimane rated consonant and dissonant chords and vocal harmonies as equally pleasant despite showing similar discrimination abilities and aesthetic responses to familiar sounds and acoustic roughness to those found in Western experimental subjects (McDermott et al., 2016). Bolivians, in turn, described some preference for consonance but to a lesser degree than US residents. The authors hypothesize that these results reflect levels of exposure to Western culture and that, therefore, a preference for consonant chords are unlikely to reflect innate biases or exposure to harmonic natural sounds.

We can perhaps posit that the percept of intervallic or harmonic consonance and dissonance is psychoacoustically active across musical styles and cultures. Degrees of dissonance can reflect dramatic intent, differing regulatory needs and goals, compositional and performative boundaries of style, as well as individual histories and cultural tradition (Dowling & Harwood, 1986). It seems an almost peculiar blind spot amongst some cognitive scientists to avoid the conclusion that high degrees of dissonance can be just as desirable as consonance. To highlight just a very few examples from traditional musical traditions, beating overtones are actively sought out in many Middle Eastern, North Indian, and Bosnian musical cultures (Vassilakis, 2005) and are a necessary condition for music to be considered “lively and full” in Balinese gamelan orchestras (see Spiller, 2004). Harmonising in a minor second relationship to a drone or even singing in “unison” a semitone apart are common techniques in North Indian ragas (Maher, 1976) or in Croatian traditional ‘ganga’ singing (Vassilakis, 2005). Scalar microtunings too are constructed in performance to appropriately enhance “pungency” in Anatolian Greek migrant communities (League, 2016). Of course extreme dissonance also characterises much Western art music and many modern-day styles. The central point to be clear on from the above discussion is that, ultimately, a generalisable account of music needs to account for the fundamental capacities and psychophysiological relevancies of perceiving and producing pitched tones in fine-grained relative relationship to others, rather than any absolute values of consonance or dissonance.

---

[60] Given the choice of approaching a visual imprinting object associated with consonant or dissonant versions of complete melodies, however, newly hatched domestic chicks preferentially approach the consonant version (Chiandetti & Vallortigara, 2011) and an infant chimpanzee, using a computerized setup, produced consonant versions of melodies for longer periods of time (Sugimoto et al., 2010). Thus other species may well be attuned to some structural consistencies related to degrees of consonance.

Is Pitch Perception in Music Special? – Evidence from Amusia

The following paragraphs discuss the view that pitch perception in music requires more fine-grained distinctions than those which are operational in non-musical communicative interactions. Very broadly speaking, impairments in fine-grained pitch perception do seem to have detrimental effects (albeit far less severe effects) beyond the domain of music. A particularly pertinent finding from recent research into individuals with congenital amusia is that the impairment typically stems specifically from a lack of conscious access to perceptions of small pitch differences.

Pitch clearly plays a critical role across the full breadth of human vocal interactions. Rising and falling patterns are consistently functional in emotionally regulating infant arousal levels across cultures (e.g. Trehub et al., 1993; Papousek, 1996); pitch areas are semantically functional in tone languages (e.g. Ooyang & Kaiser, 2015); and melodic contour can also indicate grammatical/lexical function (e.g. Burgoon et al., 2016). Albeit less well studied, more or less intentional microtunings and inflections of pitch also convey a myriad of communicative information and function in music, speech and non-linguistic interaction (e.g. Ladd, 2008). Any specificities of musical pitch across cultures are likely to reside in the stable and sustained nature of a musical tone (and corresponding correction mechanisms [see below]); the ability to match pitch relationally to external pitches (see below); and the relatively fine categorical distinctions between pitches in music. Structurally significant margins typically operate down to the level of 100cents/1 semitone or even sometimes 50cents/a quartertone (e.g. Dean et al., 2008) in music. This is considerably below clearly functional contour perimeters in speech or parent-infant communication and below distinctions between pitch areas in tone languages (see Cross et al., 2013). The crucial question then for the concerns of this thesis is whether this is reflected in some specific pitch perceptual capabilities for music or whether it is part of a generalised ability to handle fine-grained pitch structures in communicative contexts. Leading attempts to address this question have been investigations into individuals with acquired (see Clark et al., 2015) and, perhaps even more significantly, congenital amusia (Ayotte et al., 2002).

Despite some debate on the prevalence of amusia, the validity of its diagnostic tools (Henry & McAuley, 2010; 2013; Pfeifer & Hamann, 2015; Vuvan et al., 2017)[61], and the extent to which it generalises across cultures and boundaries of musical style (Cross, 2010), we can safely state that a small subset of individuals exhibit difficulties relating to our culture's music. It also seems clear that the predominant underlying impairment concerns fine-grained pitch perception (see Vuvan et al., 2015), either rooted in deficits in pitch processing (e.g. Peretz & Hyde, 2003) or pitch specific memory (Albouy et al., 2015, 2013; Gosselin et al., 2009; Williamson et al., 2010; Williamson and Stewart, 2010). Amusics are typically described as being impaired at detecting small pitch deviations both in synthetic music and in purer psychoacoustic studies. Pitch change detection tasks in which participants hear an isochronous sequence of tones, for example, reveal a diminished ability in amusics to distinguish intervals that are smaller than two semitones (200cents)[62] (e.g. Peretz et al., 2005 [70% performance at 25cents compared to ceiling performance in controls]). Two-interval forced choice paradigms provide similar (albeit slightly more fine-grained) results (Foxton et al., 2004; Liu et al., 2010; Tillmann et al., 2009). As a group, strong deficits for thresholds of minimum pitch distinctions are also statistically evident when participants are asked to perform tasks that require processing of pitch direction (Jiang et al., 2013; Liu et al., 2012; Williamson et al., 2012)[63]. This basic psychoacoustic constraint is understandably more definitively ruinous in music and has been seen to explain why this impairment can seemingly occur in relative isolation from speech disorders (e.g. Peretz et al., 2002; see Peretz, 2006). More recent reports and studies on congenital amusia have, however, revealed a more nuanced picture.

The majority of the studies referenced in the previous paragraph (also Jones et al., 2009; Omigie et al., 2013) have documented a degree of overlap between their amusic and control groups in terms of pitch thresholds in both pitch change detection and/or pitch direction judgments (see Tillmann, 2016). In other words, a subset of amusic participants exhibited thresholds within the normal range of controls. Significantly, perceptual pitch processing deficits in amusics have, in recent years, also been reasoned to impair other linguistic and interactional functions. Studies have highlighted deficits in congenital amusia (albeit less severe than those found in music) in processing functional speech intonation (question vs. statement [Patel, 2008]), pitch manipulated French syllables (Tillmann et al., 2011)[64], in decoding emotional prosody in speech (Thompson et al., 2012), and also in contrasts in lexical pitch in tone languages (Mandarin or Thai; e.g. Nguyen et al., 2009; Tillmann et al., 2011b). Crucially, from a comparative perspective, amusic speakers of Mandarin showed a similar pattern of musical impairment as Western amusics with half of the participants in a seminal study displaying impairments in the discrimination and identification of Mandarin lexical tones (Nan et al., 2010). This indicated to the authors that the pitch disorder at the root of most congenital amusia is not specific to music or culture but is rather general in nature[65]. A later study confirmed these findings, whilst additionally proposing significant correlations between pitch threshold and lexical tone perception, music perception *and* production, but not between lexical tone perception and production (Liu et al., 2016). This supports a partial dissociation between pitch perception and production (see below) and that amusia in mandarin

---

[61] Considering that almost half (41%) of individuals diagnosed as amusic describe a relatively normal experience in terms of musical engagement and appreciation, we could perhaps even challenge the use of term itself (Omigie et al., 2012).

[62] In contrast to controls who generally perform at ceiling for intervals as low as a quarter of a semitone (25 cents) (Hyde & Peretz, 2004).

[63] In accordance with this research Patel (2008) proposed a "melodic contour deafness hypothesis" for amusia (see also Liu et al., 2012).

[64] A particularly curious finding in this study was that the amusic group performed better for syllables than tones at the smallest pitch changes (25 cents). They even achieved normal performance for syllables with intermediate pitch changes (100 cents/1semitone – an interval they struggled with in musical contexts).

[65] A genetic basis for pitch processing deficits in amusia is also evident from studies showing strong heretability (Peretz et al., 2007 [70-80%]) and greater similarity in scores of recognising melodic anomalies in monozygotic twins than dizygotic twins (Drayna et al., 2001; Seesjaarvi et al., 2016).

speakers is “characterised by severely impaired music perception and production, mildly impaired speech perception, and largely intact speech production” (ibid, p. 563).

While psychoacoustic limitations and deficits of memory representation in pitch processing appear as the behavioural correlates of amusia, the neurologically functional root appears to go deeper. Fundamentally, sensory information seems to appear normally in the brain of amusics. Lehmann and colleagues (2015) documented gradations of auditory brainstem responses correlating with musical experience/ability to fine-grained pitch stimuli. However, Peretz (2016) reasons that this may reflect reduced transmission of cortical efferences (Coffey et al., 2016), and otherwise we find no irregularities in cochlea (Cousineau et al., 2015) or midbrain functioning (Liu et al., 2015). Furthermore, amusic individuals with a substantial pitch perception deficit demonstrate clusters of pitch-responsive voxels in the primary motor cortex (A1) that are comparable in extent, selectivity, and anatomical location to those of control participants (Norman-Haignere et al., 2016).

A number of varied neuroscientific studies have now also found surprisingly normal brain activity in response to pitch deviations that often go unnoticed by amusic subjects. This suggests that “the core deficit in congenital amusia resides in a lack of conscious access to processed pitch deviances” (Peretz, 2016 p. 859). Peretz and colleagues (2009) presented a seminal study in which amusic participants demonstrated normal early right anterior negativity (ERAN) to melodic deviations (cf. Koelsch et al., 2011). FMRI blood oxygen level-dependent (BOLD) responses to successively patterned discrete pure-tones ranging from 25 cents to 2 semitones were also similar in amusics to those from control participants (Hyde et al., 2010). Both showed a positive linear correlation between BOLD responses in bilateral auditory cortices and increases in pitch distance. A study of mismatch negativity (MMN) brain responses to deviants in repeated tone sequences also produced congruous results (Moreau et al., 2013). In small deviations (e.g. 25 cents) amusics, as usual, reported poor detection but showed normal MMN brain responses. Significantly, in amusics the smaller deviations, in contrast to their responses to larger pitch changes, were not followed by a typical positivity (P300/P600) associated with conscious recognition. Finally a more recent ERAN study (Zendel et al., 2015), on violations of expected tonality, demonstrated an evoked ERAN and a similar interference effect in amusics and controls in a click-detection task. Despite being instructed to ignore pitch violations both groups exhibited decreased scores for click-detection accuracy – presumably due to an auditory attentional “blink” (Raymond et al., 1992). However, when explicitly focused on pitch-detection, an evoked ERAN and P600 was only evidenced in the control participants. This study is especially significant in suggesting that tonal knowledge can be implicitly acquired in amusics, albeit remaining not consciously accessible[66]. The capacity to engage appropriately with a musical tone appears, uniquely to some extent, to be contigent upon fine-grained pitch discernments *and* conscious access to perceptual information. The latter may relate to degrees of volitional control in music-specific pitch correction mechanisms (discussed below).

The Psychophysiology of Pitch Production

The following four sections will follow a similar structure to the preceding discussion on the psychophysiology of pitch perception with regard to vocal pitch production. Fundamentally there appears to be a greater degree of dissociation between perception and production than was the case with pulse. I will present a comparative analysis of the volitional control of individual parts of the vocal mechanism and will argue that it is particular psychological/attentional correction mechanisms inherent in “sensorimotor harmonisation” that are most crucial in distinguishing the production of a specifically musical tone.

In exploring the production of pitched tones it makes sense, especially from an evolutionary perspective, to focus on vocal capabilities. The most fundamental features of our vocal mechanism are shared with other species and are commonly operational in pre-linguistic communication, linguistic speech, and music. It can therefore be subjected to constructive comparative analysis. Furthermore, singing and percussive rhythmic actions are surely the modes by which music was primarily enacted by our evolutionary ancestors. They remain notably dominant in modern-day hunter gatherer societies (see Morley, 2013). Some modern-day pedagogies (e.g., Kodály) even start from the premise that the human voice is the natural musical instrument. One could, therefore, perhaps regard melodic instruments as being, first and foremost, an extension of the voice. That said, we should perhaps be careful in presuming this somewhat poetic notion too absolutely. There may be some important psychophysiological differences between the manifest sonic feedback from an “external” instrument as opposed to that from one’s own “internal” voice (see below).

Evolutionary and comparative perspectives on vocal capabilities have typically concentrated on evidence for the range of possible vocal sounds and degrees of volitional control in humans, with theoretical discussions focused on the relevance of findings to both language and music. The next few sections follow some of the principal research in these areas. However, in keeping with the central focus of this thesis, I will try to highlight root capabilities that may correlate to distinctions of a “simple” tone being musical. As we shall see, even the production of a single sustained fundamental frequency is not trivial. It necessarily requires a consistency of muscular tonus and is reliant upon an ability to generate volitionally sound that varies independently in loudness, pitch (F0), and timbre. Significantly, this level of independence of acoustic dimensions in production appears gradually in human infant development (Wermke & Mende, 2009) and is not fully shared with other

---

[66] Another way of describing this is that the abnormalities found in amusia seem to stem from a disturbed top-down feedback from the inferior temporal gyrus (IFG) to the superior temporal gyrus (STG). Peretz (2016) suggests that this functionally connects to conscious access/experience, in that the IFG is often operational in amplifying and refining auditory processing in the STG (Optiz et al., 2002), and because cognitive amplification is a crucial component for both predictive coding (Garrido et al., 2009) and consciousness (Deheane & Changeux, 2011).

primates who exhibit a relatively close linear relationship in vocalization between F0 and intensity with changes to subglottal pressure (Demolin & Delvaux, 2006). Many bird and mammalian species use modulations of spectral shape and pitch for communicative purposes and in distinguishing functional intent (e.g., Elie & Theunissen, 2016; Fitch, 2005; Pisanski, et al., 2016; Morton, 1977). However, as noted by Patel (2017) this does not prove that they are modulated independently in sequences to convey distinct types of information. It is notable, therefore, that even voicing a single musical note is contingent upon complex interactions and degrees of independent control of the muscles involved in breathing and vocal production: Glottal adductor and abductor muscles serve to open and close the vocal folds that regulate levels of subglottal pressure thus affecting F0 and intensity; pharyngeal constrictor muscles control the width of the pharynx hence altering the relationship of F0 and higher harmonic partials (i.e., timbral correlates); extrinsic laryngeal muscles control the vertical positioning of the larynx in the throat; and intrinsic laryngeal muscles modulate length, stiffness, and thickness of the vocal folds through contractions of the cricothyroid and thyroarytenoid muscles and, crucially, allow an uncoupling of intensity and F0 (Sundberg & Rossing, 1990; Titze, 1994).

Given the primacy of the voice, it would be tempting to assume some form of physiological and neurological coupling of pitch perception and embodied vocalisation (such as we found between pulse and movement) – some form of "inner singing" (e.g. Janata, 2001; Kalakoski, 2001). Indeed areas of the Rolandic operculum, in particular those areas that are likely to represent laryngeal and pharyngeal articulation, have been shown to be active during the emotional processing of instrumental musical stimuli (Koelsch et al., 2006). Belyck and Brown (2017) also hint at some involvement of the larynx motor cortex in the perception of pitch and Brown and Martinez (2007) describe activations of premotor vocal areas during melody and harmony discrimination tasks. However, there does seem to be, also, a considerable degree of dissociation. Correlations between a variety of singing accuracy and pitch perception tasks have been evidenced in some studies (e.g., Amir et al., 2003; Watts et al., 2005; Moore et al., 2007; Estis et al., 2009, 2011) and contradicted in others (Bradshaw & McHenry, 2005; Dalla Bella et al., 2007; Pfordresher & Brown, 2007; Moore et al., 2008). We should perhaps also be mindful of the point that we are generally more forgiving of degrees of deviations in tuning in singing than instrumental music across cultures (Arom et al., 2007; Pfordresher & Brown, 2017). Hutchins et al., (2012) revealed judgments of sung melodic intervals being "in tune" up to a threshold of roughly 60 cents – a difference beyond the boundary between discrete scalar steps[67]. The evidence from studies of congenital amusics is also equivocal. Unsurprisingly congenital amusics are consistently rated as poor singers (Ayotte et al., 2002). As a group they make considerably more pitch errors in singing popular songs (Dalla Bella et al., 2009) or matching single pitches (Hutchins et al., 2010) than controls. However, we also find individual amusics whose singing is much better than their deficits in perceptual ability would suggest (Hutchins et al., 2010) and improvements are evident in amusics as a whole when directly imitating a model (Tremblay-Champoux et al., 2010). In an interesting study comparing performance on a synthesized slider versus singing, Hutchins and Peretz (2012) sought to uncover the cause of poor pitch singing. Results showed that about 20% of singers had a vocal-motor control impairment, 35% had a sensorimotor deficit (related to timbre), and only 5% had a perceptual deficit. Perhaps the most accommodating theoretical model for this and the current state of evidence is the 'linked dual representation model' proposed by Hutchins and Moreno (2013). This predicts that vocal information can be encoded *distinctly* as "a symbolic representation, such that we gain conscious knowledge of the identifiable features of the vocal stimulus… (and) as a motoric representation, such that it enables reproduction, imitation, or generative production" (p. 7).

Crucially lacking in these discussions is an experimental understanding of how we merge perception and production in harmonizing our output synchronously with others. The following few sections will, therefore, move towards suggestions for a new experimental paradigm to explore 'sensorimotor harmonisation' and some corresponding hypotheses which I hope to explore in future postdoctoral work (see below).

The Evolution of Vocal Capabilities

There have, as will be discussed below, been substantial evolutionary changes in hominin vocal physiology with considerable implications for musical and linguistic capabilities. Critically, however, evidence for increased vocal flexibility and fine control of sustained air pressure cannot account entirely for the *specific* features and capabilities inherent in the production of a musical tone.

Physiological constraints

Much of the evolutionary literature on vocal physiology has been framed with reference to discussions on the capacities for speech (e.g. Fitch, 2000; Johansson, 2005). However, evidence for phylogenetic developments of vocal physiology is equally relevant to music (Frayer & Nicolay, 2000), in particular singing. In fact it is perhaps best described in terms of fundamental physiological features required for controlled and varied vocalizations be it linguistic, musical, or otherwise (Morley, 2002; see Morley 2013). It is also worth pointing out that evidence for physiological capacities does not imply that the capacities were used in musical or speech-like context. We can, however, aim to infer some comparative hypotheses and a time before which such behaviours were not available in our evolutionary ancestors.

---

[67] Notably if we use microtuning in human singing rather than instrumental music it counteracts the claim by Araya-Salas (2012) that nightingale wren's song cannot be considered a suitable analogy to human music on the basis of intervallic consistency being disproportionately variable.

The most crucial physiological features involved in vocal production – the larynx, tongue and supralaryngeal tract - consist of cartilaginous and soft tissue and hence are not preserved in the paleontological record. However, some researchers have inferred variously from the mandible (jaw), basicranium (bottom of the skull), hypoglossal canal (Kay, et al. 1998)[68], and on rare occasion the hyoid bone (e.g. Arensberg et al. 1990). Following a seminal paper by Lieberman and colleagues, it was widely thought that the vertical positioning of the larynx inferred from degrees of basicranial flexion – the degree of curvature of the underside of the base of the skull – was indicative of increasing vocal versatility. The argument was that in other primates a high larynx restricts the resonance, and consequent palette of timbres, produced by the pharyngeal cavity (Lieberman, 1984; Budil, 1994; Duchin, 1990; Laitman & Heimbuch, 1982; Laitman & Reidenberg, 1988). In contrast the human larynx descends in the course of ontogeny (Magriples & Laitman, 1987) thus connecting the upper respiratory and digestive tracts (Laitman, 1984) and is permanently positioned low in the throat[69]. This appears to have emerged gradually in human evolution. Laitman and colleagues' (1979) examination of KNM-ER 406 (*Paranthropus boisei*) and Sterkfontein V (*Australopithecus africanus*) both show no basicranium flexion. Evidence for basicranial flexion first appears with KNM-ER 3733 (*Homo ergaster*) at 1.75 MYA (Ibid. 1979) and fully modern basicranium is first fully displayed in *Homo heidelbergensis* 300-400 thousand years ago. Budil's (1994) supralaryngeal reconstructions also concur with a view that a modern-day upper respiratory tract first appeared in *Homo heidelbergensis* and that *Australopithecus* retained ape-like characteristics.

However, the inference of a constraint on vocal versatility due to a high resting larynx position has proved problematic. It seems likely that the primary selective influence on basicranium flexion was increasing brain size rather than communicative versatility (Spoor & Zonneveld, 1998; Ross & Ravosa, 1993). Furthermore, in fully upright hominins such as Neanderthals a lack of basicranial flexion is less significant with regard to supralaryngeal volume and variety of vocal production (Morley, 2003). Modern-day comparative evidence also contradicts the inferred constraint. Morley (ibid.) notes, for example that individuals with Down's syndrome appear not to be impaired by a high laryngeal position in terms of being able to produce a full range of sounds (Scothern, 1992). Even clearer is a recent study on living macaques using x-ray videos to quantify vocal tract dynamics during vocalization, facial displays, and feeding (Fitch et al., 2016). They argue that the macaque could easily produce an adequate range of speech sounds to support spoken language. This conclusion has since been supported by Lieberman (2017; see Fitch et al., 2017) and has been taken to indicate, as hypothesised by Darwin, that the evolution of human speech capabilities required neural changes rather than modifications of vocal anatomy. These conclusions can equally be applied with respect to music/singing. From our current perspective it is also worth highlighting that, while putative comparative distinctions between human and primate vocal mechanisms may affect the types of sounds that can or could potentially be manifest as music, it is the sustained nature of a tone that arguably makes it musical rather than degrees of versatility. Of potential relevance, therefore, is a rarely mentioned physiological comparative study by Han et al., (1999) who show that the thyroarytenoid muscles in humans uniquely exhibit rare slow tonic muscle fibers that, unlike most muscles, do not twitch and sustain prolonged, stable, precisely controlled and fatigue resistant contractions.

In general it seems fair to conclude, from the above discussion, that the evolution of our capacity to produce a sustained tone was not contingent upon changes to the fundamental vocal apparatus. A possible exception to this is the necessary breath control to produce a sustained tone. Paleoanthropological discussions on breath control have again been primarily framed around language. However, a fine control of sustained air pressure is clearly also a crucial feature in singing. Given that discrete pitches in music are, by definition, relatively extended utterances, it may be even more fundamental to music than language. It is, most crucially, an inherent and constituent part of our capacity to produce even a simple musical tone. Some researchers have suggested that the control of breath was already enhanced, compared to extant primates, in *Homo ergaster* (Frayer & Nicolay, 2000). However, the evidence remains equivocal and contrasting interpretations regarding the emergence and evolutionary trajectory of breath control in the hominin lineage have been presented in the literature[70].

---

[68] Kay and colleagues (1998) suggested that the relative size of the hypoglossal canal reflects levels of innervation and hence fine control of the tongue. They present analysis that a shift from within the range of modern *Pan troglodytes* and *Gorilla gorilla* to the range of modern humans occurred between *Homo habilis* and *Homo heidelbergensis*. However these findings were undermined by De Gusta et al. (1999) who, with a larger data set, found that many non-human primates, as well as some *A. africanus* (A.L.333-45, A.L.333-105, A.L.333-114) and *A. boisei* (Omo L.338-y-6), have/had hypoglossal canal dimensions within the range of modern humans. Additionally, an analysis of the hypoglossal canal volume, and the number (and size) of axons in the hypoglossal nerve, in five modern human corpses revealed no statistically significant results (Ibid.).

[69] According to Nishimura humans have undergone a two-part descent of the larynx. The first of which, the descent of the cartilaginous skeleton of the larynx relative to the hyoid bone, is they argue shared with chimpanzees but not monkeys (Nishimura et al., 2006; 2008).

[70] MacLarnon and Hewitt (1999; 2004) suggest that evidence from comparison of the vertebral canals of fossil hominids and those of existing primates suggests a major increase in thoracic innervation in the course of hominid evolution. This, they argue, provided enhanced control of the intercostal muscles and control of subglottal pressures. Comparing human, fossil hominids, and non-human primates, they note that vertebral canal dimensions correlate with dimensions of the spinal cord and that spinal cord weight and body weight generally relate logarithmically. The one exception is the thoracic region. Although thoracic spinal nerves serve many functions, MacLarnon and Hewitt (1999; 2004) argue that the most likely cause of the observed increase is enhanced breathing control (see MacLarnon and Hewitt 2004 p. 192-3 for full details). Furthermore, their analysis of fossil hominins suggests that thoracic regions of *Australopithecines* (AL288-1, Sts 14, Stw 431) and *Homo ergaster* (KNM-WT 15000) were, size-adjusted, within the range of extant primates. Neanderthals (La Chapelle, Shanidar 2, Shanidar 3, and Kebara 2) and early modern humans (Skhul 4), on the other hand, have dimensions within the range of modern humans. The case of KNM-WT 15000 ('the Nariokotome boy') has been the subject of some interesting and instructive debate. Frayer and Nicolay (2000) argue *contra* MacLarnon (1993) and Walker (1993) that the thoracic nerves are not the most crucial for fine breath control but rather that "most of the muscles involved in speck respiration are innervated by either cranial or cervical nerves" (p.226). They continue that, as the only vertebra preserved from KNM-WT 15000 is within modern human range, this suggests that *Homo ergaster* may have had modern-human equivalent control of breath. A complete perspective integrating these two positions is taken by Iain Morley

The Neurology of Vocal Control

The principal evolutionary distinctions between vocal capabilities in humans and other primates are likely to have been driven by neurological changes. The following sections therefore offer a review and comparative analysis of the neurology of vocal control in human and non-human primates. Behavioural studies confirm degrees of independent voluntary control over the articulatory mechanisms, subglottal pressure, and the laryngeal muscles in primate species. However, the levels of voluntary control are qualitatively far removed from those exhibited in humans.

Some paleo-neuroanatomical inferences regarding the neurology of vocal control have been suggested from available archaeological endocasts. The exact degree to which one may reason brain function from the physiology determined from endocasts is debatable (e.g. Leakey, 1994) so we cannot perhaps expect to reveal anything with certainty. However, from about 2 million years ago there does seem to have been disproportionate growth in Broca's area starting with *Homo habilis/Homo rudolfensis* (KNM-ER 1470) (Holloway, 1983)[71] and a further increase, judging by growth in the lateral tuba corresponding to Broca's area, by the time of *Homo ergaster* (KNM-WT 15000) (Kochetkova, 1978). This has been argued to relate to emerging linguistic capabilities. However, PET studies have shown that the classic Broca's area (Brodmann's areas 44 and 45) is operational in motor functions and, in particular, motor planning[72] (Peterson et al., 1988). Brown et al. (2006) have supported this by showing BA 44 and 45 activation in improvising both sentences and melodic phrases. Additionally, the analogous area in macaque monkeys is implicated in orofacial musculature control (Petrides et al. 2005). So it seems likely the growth of Broca's area indicates an increased capacity to plan and execute complex vocal patterns (which may well initially have been neither linguistic nor musical).

More illuminating evidence of the evolutionary adaptations that have occurred with respect to the neurology of vocal production in humans comes from comparisons between modern-day primates and humans. Here the key distinction, both behaviourally and neurologically, appears to be between instinctive utterance and volitional control. Different authors have preferred different ways of describing and framing this distinction. Hage and Nieder (2016), for example, structure their recent review on a dual neural network model for the evolution of speech and language in terms of a 'primary vocal motor network' (PVMN) and a 'volitional articulatory motor network' (VAMN). Offering a slightly different angle, but addressing many of the same concerns, Owren and colleagues (2011) instead describe a distinction between 'production-first' and 'reception-first' development in human and nonhuman primate vocalization – the latter being marked particularly as a form of vocal learning. Finally another recent review paper by Gruber and Grandjean (2016) classifies emotional vocalizations in primates according to four dimensions of learning, control, emotion, and meaning. These, they argue, can combine to form three main types of vocalization: "(a) spontaneous and uncontrolled, (b) spontaneous but controllable via voluntary inhibition or amplification of the call, or (c) produced intentionally by the signaler with the intent to induce an emotional state or a specific response in the audience" (p.187).

The neurological pathways involved in involuntary emotional vocalization in primates are reasonably well understood (see Jurgens; 2002; 2009). Of primary importance is a network incorporating the periaqueductal grey matter (PAG) of the midbrain and the anterior cingulate cortex (ACC). Both of these areas produce vocalizations when stimulated electrically or pharmacologically (e.g. Dujardin and Jürgens, 2005). In particular, control of laryngeal muscles is thought to involve the neurological pathway from the PAG via the nucleus retroambiguus to monosynaptic connections to the laryngeal motoneurons in the nucleus ambiguous (e.g. Vanderhorst et al., 2001). A second neural pathway includes the primary motor cortex and two subcortical loops that modulate vocal motor commands from M1 and subsequently send modified motor programs via the ventrolateral thalamus back to M1. This area of research is reviewed more extensively elsewhere (e.g.

(2003) who insightfully argues that, in fact, both positions are valid and complementary. His argument and conclusion is worth quoting at length here as it relates directly to capabilities involved in producing a simple sustained and musical tone:

> Whilst the thoracic intercostal nerves are indeed responsible for some control of breathing (Palastanga et al., 2002/[2006]), they seem to be implicated in additional fine control of a function that is principally controlled through cervical innervation. Cases of spinal injury in modern humans confirm this. Spinal injury below C6 still allows control of breathing without assistance, and vocalization capabilities indistinguishable from normal, except that the patient is limited in the depth of inhalation. What the thoracic nerves seem to allow is control of especially prolonged vocalizations. The duration of vocalizations can be affected in patients with spinal injuries below the cervical vertebrae, but the quality of such vocalisations is not affected by the loss of thoracic control (The PoinTIS Spinal Cord Occupational Therapy site of the SCI Manuals for Providers; pers. comm. G. C. Morley, clinical physiotherapist). (p. 92-93)

He goes on to infer that:

> [i]f the cervical vertebrae of KNM-WT 15000 are indeed of modern dimensions as Frayer and Nicolay (2000) and MacLarnon (1993) assert, this indicates that there is no reason why we should not attribute to *Homo ergaster* the ability to voluntarily moderate its breathing sufficiently well to produce vocalisations controlled for pitch, intensity and contour. Limiting the variation in pitch and contour would have been the degree of development of the upper respiratory tract and, in particular, the larynx… What it may have lacked was the ability to produce extended voluntarily controlled vocalisations. It seems that this is an ability which has developed over the course of human evolution, indicated by thoracic canal innervation. (Ibid, p.93)

[71] This skull from East Turkana is now more commonly classified as *Homo rudolfensis* which is not a direct descendant in the human lineage (Johanson & Edgar, 1996)

[72] It was instead Brodmann's area 47 that was implicated in semantic associations.

Jurgens, 2009; Gruber & Grandjean, 2017; Ackermann et al., 2014)[73]. The most crucial conclusions for the principal concerns of this thesis are that models of non-human primate vocalisation (e.g. squirrel monkeys) appear to provide a suitable basis also for basic reflexive emotional utterances in humans (e.g. crying, groaning and/or laughing) (Wild et al., 2003; Fitch, 2006; Brudzynski, 2009; Bryant and Aktipis, 2014), and also that these pathways remain operational (albeit not exclusively so) in more complex volitionally controlled vocalization (e.g. Schulze et al., 2005)[74]. Functionally, the most important identified difference between humans and other primates is that in humans there is a direct connection between M1 and the laryngeal motoneurons in the nucleus ambiguus (Jurgens, 1976; 1994)[75]. Behaviourally, this may be reflected in a lack of ability in nonhuman primates to volitionally decouple laryngeal motor activity from species-typical audiovisual displays (e.g. Knight, 1999; see below) and has been argued to support efficient voluntary control of vocalization in humans (Fischer & Hammerschmidt, 2011; Fitch et al., 2010; Fitch, 2011; Jarvis, 2004; Simonyan & Horwitz, 2011)[76].

Vocal Control in Nonhuman Primates

Vocalisations by non-human primates are often thought to be predominantly innate and stereotyped calls, uttered almost exclusively in direct correlation of affective state (Hammerschmidt & Fischer, 2008 [in Oller & Griebel Eds.]; Jurgens, 2002). Congenitally deaf squirrel monkeys or those raised without auditory models, for example, develop a full species-typical repertoire of calls and use them in appropriate motivational and social contexts (Winter et al., 1973, Hammerschmidt et al., 2001). Unlike remarkable evidenced capacities in some avian species, pinnipeds, cetaceans, bats, and elephants (Janik and Slater, 1997; Catchpole and Slater, 2003; Ridgway et al., 2012; Stoeger et al., 2012), vocal learning and imitation have traditionally been assumed to be absent or at best highly restricted in great apes (Fitch et al., 2010; Lameira 2014). It has subsequently been supposed that human vocal capacities have no direct ancestor in nonhuman species (Bolhuis & Wynne, 2009).

However, increasing degrees of vocal control have been evidenced in primate studies. Various species have, in operant conditioning tasks, been shown to be able to produce or inhibit vocalisation (Coude et al., 2011, Hage & Nieder, 2013, Sutton et al., 1973, Aitken & Wilson, 1979; Hihara et al., 2003; Koda et al., 2007; Sutton et al., 1985). Wild chimpanzees, for example, can voluntarily inhibit food grunts (Wilson et al., 2007), alarm calls (Cheney & Seyfarth, 1985), greetings (Laporte & Zuberbuhler, 2010), and responses to unknown conspecifics (Wilson et al., 2001)[77]. Many field studies also demonstrate non-human primates altering their species-typical vocalisations depending, for example, on the social context or presence of a predator (Seyfarth et al., 1980); the awareness or behaviour of others (Crockford et al., 2012; Wich & De Vries, 2006); other particular external events (Quattara et al., 2009); the nature of visual cues (Hage et al., 2013); or even intentional goals (Gruber & Zuberhuhler 2013; Townsend et al., 2014). Furthermore, detachment between produced acoustic signal and physiological correlate has been observed in wild orangutans (Wich et al., 2012), as has a degree of functional flexibility of similar calls in wild bonobos (Clay et al., 2015; Genty et al., 2014). Finally, a number of individual-specific and population-specific voiced calls in great apes that do not conform to genetic and ecological divergence (e.g. Lameira et al., 2015) suggest to Lameira and colleagues (2016) that “vocal fold control may play indeed an active role in shaping the composition of the voiced repertoire of great apes” (p. 2).

Hage and Nieder (2016) summarise that “results indicate that monkeys are capable of some types of vocal learning, in other words they are able to cognitively control the onset of their vocal output and to modify their vocal patterns within the range of their natural repertoire” (p. 814). More detail and further references can be absorbed from that paper (see also Gruber & Grandjean, 2016). For the purposes of this thesis, the more directly pertinent question is whether nonhuman primates can independently manipulate breath, supralaryngeal, articulatory and laryngeal mechanisms in vocal production? More

---

[73] Zarate (2013) summarises neatly that “in humans, these networks form a tripartite hierarchy of vocal motor control (see Simonyan and Horwitz, 2011): (1) the reticular formation constitutes the lowest level at which complete vocal patterns are generated; (2) the next level is comprised of the ACC and the PAG, which are attributed with the voluntary initiation and emotional/motivational control of vocalizations (Jürgens, 2002, 2009); and (3) the highest level of vocal control occurs in M1 (and its modulatory brain regions), which is associated with the generation of learned vocalizations, such as speech and song (Jürgens, 2002, 2009)” (p. 2).

[74] It is important to note that so far we are describing a general functional network for human vocalization rather than attempting to highlight any specificities of music. A range of tasks that have been confirmed to recruit many regions within vocal motor and sensory networks are selected in Zarate (2013) - “word or letter generation (Paus et al., 1993); syllable repetition (Riecker et al., 2005); singing a note repeatedly (Perry et al., 1999), in a sustained fashion (Zarate and Zatorre, 2008), or while changing vowels in particular rhythms (Jungblut et al., 2012); repeating syllables, spoken words, and sung or hummed melodies (Özdemir et al., 2006); humming, speaking, or singing lyrics of a well-known song (Formby et al., 1989; Jeffries et al., 2003); reciting the months of the year or singing a familiar melody (Riecker et al., 2000); telling a story (Schulz et al., 2005); improvising word phrases, melodies, or harmonies (Brown et al., 2004, 2006); spontaneous and synchronized speaking and singing (Saito et al., 2006); and singing an Italian aria (Kleber et al., 2007)” (p.4). It seems likely therefore that all these capacities are built exaptively upon existing systems for vocally communicating emotional state and arousal with a heightened level of voluntary control in humans.

[75] In contrast with chimpanzees we also see a distinct direct monosynaptic connection between the primary motor cortex and the phonatory, articulatory and respiratory areas of the medulla in the brain stem (Okanoya & Merker, 2007).

[76] In a recent review article on the ‘vocal brain’ Belyck and Brown (2017) however argue that, although this connection can perhaps account for increased volitional control over the laryngeal muscles, it does not seem sufficient to account for the novel engagement of the respiratory musculature seen with LMC (Larynx Motor Cortex; see Brown et al., 2007) stimulation in humans. They suggest that the human motor cortex has been evolutionarily reorganized to bring the three major components of vocalization − expiration, phonation, and articulation − into close proximity, thus creating a particularly efficient interaction - a “small-world architecture” (e.g. Sporns, 2006; Sporns and Zwi, 2004).

[77] The prefrontal cortex is thought to be of primary importance (Owren et al., 2011) correlating with its supposed role in human action inhibition (e.g. Ridderinkhof et al., 2004).

precisely can they volitionally isolate handling of fundamental frequency? Evidence collected in the last decade, predominantly from captive or enculturated great apes, would appear to indicate at least a partial affirmation on all of the above. A latent capacity to control airflow and vocal fold adduction is evidenced in novel attention-seeking grunts in captive chimpanzees (Hopkins et al., 2007). Unvoiced articulator manipulation in orangutans has been argued to provide some potential homology to linguistic consonants in humans (Lameira et al., 2014) and some recent studies on enculturated great apes would appear to suggest latent capacities to manipulate the vocal system voluntarily and thus manipulate both F0 and acoustic formants (Taglialatela et al., 2003; Perlman & Clark, 2015; Lameira et al., 2015). Captive orangutans have also been reported to imitate whistles produced by humans (Wich et al., 2009) and will socially and spontaneously match single, duple, and/or triple whistles correctly (Lameira et al., 2013). Individuals produce these whistles in different ways indicating some independent supralaryngeal control over the muscle systems of the upper lip, lower lip and respiration. However, nothing in the results indicates an attempt to match the F0 of the human whistle.

In a more recent study (Lameira et al., 2016) the same authors directly challenged the assumption that great apes lacked direct voluntary control of the larynx. The captive orangutans had developed an idiosyncratic - ingressively voiced - “wookie” vocalization which the authors argued to exhibit a unique spectral profile among the orangutan vocal repertoire. Intriguingly, one individual – Rocky - would spontaneously match human “wookies” with adjustments to voice frequency beyond the typical range of unprompted versions. This is a critical finding both in terms of the implications of laryngeal control and also the relevance of frequency “matching” in an interactive setting. However, we should note that, although these are not species-typical vocalisations, the individuals had been producing them for over 6 years with an unknown origin. Furthermore, Rocky did not so much match the frequency of the human version as move preferentially in that direction. Finally the degree of direct vocal control of the larynx is ameliorated by correlations between frequency and loudness and length of the call. Thus, part of the manipulations of frequency occurred via control of the breath and subglottal pressure[78]. This is also common in songbirds (see Amador & Margoliash, 2013) and is a characteristic of most human vocalization (Pabon et al., 2011). However, we can posit that the acoustic separation of frequency from loudness and intensity is most pronounced in music and is reflected in a biologically uncommon disconnect between pitch and arousal ratings (Ilie & Thompson, 2006; Gussenhoven, 2002; see chapter 4).

Clearly degrees of vocal control in humans have evolved to levels that are highly exceptional and perhaps even magnitudes removed from anything found in other primates. However, our understanding and interpretations of capacities in nonhuman primates seems to be expanding and we can certainly already advocate that latent modern-day capacities suggest a neuro-behavioural foundation for vocal control (and hence some crucial preconditions for the production of controlled speech and a sustained musical tone) in an ancient common ancestor[79`80].

Is Pitch Production in Music Special? - Evidence from Frequency-Shift Auditory Feedback Studies

In this section I argue that, although some correction mechanisms are commonly shared with speech, specific corrections are operational in maintained pitch-matching in music. Similarly to period correction mechanisms in a musical pulse, these are putatively neurologically distinct, subject to manipulations of intention, and incur awareness. However, as the evidence currently is not conclusive I will also be proposing (further along in this chapter) a new experimental paradigm for investigating ‘sensorimotor harmonisation’ which I propose to pursue in future work.

A crucial part of intentional vocalization is necessarily a capacity to correct for unexpected perturbations that may result, for example, from naturally occurring degrees of motor variance. Just as the relative stability of pulse in music and its putative specificity can be explored at a deeper lever of phase and period corrections (chapter 2), it is tempting to speculate that the relatively sustained and discrete nature of a musical tone is also foundationally characterized by specific correction mechanisms[81]. Correction mechanisms in engagement with musical pitch are as yet not well understood. In particular, an experimental paradigm studying our musical capability to match external frequencies is almost entirely lacking in the literature (see below). However, investigations with the voice feedback frequency shift paradigm on auditory control of voice F0 do provide some insight, avenues for inference and a structure for further hypotheses.

Two of the most primary and consistent findings in studies involving continuous vocalization are evidence for opposing responses occurring with a short latency of 100-150 ms (Burnett et al., 1998) and subsequent following responses with a longer latency of 250-600 ms (Larson, 1998). The presence of two responses was confirmed by Hain and colleagues (2000), who found that the direction of the second response, but not of the first response, can be modified by instruction. According to Donath et al., (2002) “the first response therefore indicates a negative feedback system stabilizing voice F0 automatically, while the second response may reflect a voluntary mechanism, which adjusts voice F0 to match an (supposed) external

---

[78] Nevertheless, the authors highlight that “different wookies produced by Rocky with equal frequencies exhibited wide differences in acoustic power, and vice versa”… and that “these observations would have been theoretically impossible if Rocky had not exercised some degree of direct control over vocal fold oscillation, and instead had only resorted to abdominal action to produce modulations at the level of vocal fold oscillation” (p.7).

[79] In the case of orangutans this could putatively suggest capacities present in a common ancestor 10mya.

[80] Pisanski et al (2016) suggest, for example, that early capacities for voice modulation may have their origin in biologically relevant cues such as sex and size-exaggeration.

[81] Although this section focuses comparatively on human interactive contexts, we should note also some research showing that pitch in adult birdsongs is also maintained through correction. In a study by Amador and Margoliash (2013) perturbed F0 auditory feedback in Bengalese finches was fully compensated for. Output returned to baseline after the perturbation was removed, thus highlighting that the sensory target is a key priority in this species.

reference" (p. 1587). Pertinently, studies of speech (Natke & Kalveram, 2001; Donath et al., 2002) have only shown opposing responses to frequency-shifted auditory feedback occurring with a latency of about 160 ms. These responses typically compensate in the region of 25-50 cents only for changes of 100cents or above (see also Liu and Larson, 2007; Jones and Keough, 2008), suggesting perhaps that they are concerned merely with stabilizing naturally occurring variation in vocal motor control. "Speech" in these early studies was prosodically uttered nonsense words (e.g. 'ta'tataa') and the delay in the corrective responses indicated that these regulate F0 at a supra-segmental prosodic level rather than on individually occurring voiced features within syllables. In a follow-up study directly comparing "speaking" and "singing" (a pitched and sustained 'ah' vowel) it was found that this initial opposing response in the singing condition was stronger (roughly 60cents) (Natke et al., 2003)[82,83]. This was thought to relate to the notion that in singing we have an internal reference F0 even in the absence of external influence. However, we should also note that when pitch-shifted stimuli are presented during continuous speech in Mandarin and English, the initial response magnitudes are in the same range as sustained vowel sounds (50–80 cents) (Xu et al., 2004; Chen et al., 2007). In singing a scale, where presumably the internal reference targets are entirely clear, most singers were reported to compensate fully for the perturbation to auditory feedback (Burnett et al., 1997; Parlitz and Bangert, 1998).

The two responses have been suggested to be manifest neurologically in recruitment of the midbrain PAG for the early more automatic pitch-shift response, with areas typically involved in actions under voluntary control (e.g. auditory cortex and ACC) operating in the later response (see Zarate, 2013). Due to FRMI temporal limitations activation of the PAG has not been determined (Zarate et al., 2010). However, Zarate and Zatorre (2008) did find activations of the intraparietal sulcus (IPS) and dorsal premotor cortex in a pitch-shifted singing task compared to normal singing. This, they suggest, is linked to the role of the IPS in transformations of sensory input for motor preparation (Astafiev et al., 2003; Grefkes et al., 2004; Tanabe et al., 2005). In contrast to these studies, Toyomura et al (2007) instructed participants to keep the pitch of the feedback voice constant as it was randomly manipulated in either direction. In this task significant BOLD activations were not evident in the ACC, rostral cingulate zone, putamen, and superior temporal sulcus (STS). Thus the corresponding differences in the task and neurological signatures to those found by Zarate and colleagues appear to reveal that these brain regions are specifically concerned with voluntary compensation responses.

In an article published in 2009 I reasoned from the above literature and results that, in speaking, only the first involuntary response occurs. I suggested that a music-specific volitionally controlled mechanism (somewhat analogous to period corrections to musical pulse) is operational in engagements with musical pitch (Bispham, 2009). A following response is, of course, in a sense what one would expect if participants hear their voice partly as if it were an external musical source to match or harmonize with. Furthermore, the slower response is necessarily contingent upon a sustained stable pitch production and a relational analysis of what has come before in comparison to the current production – in other words foundational features of a musical tone. It is also worth noting that as we approach aspects that are perhaps specific to music[84], or at least more pronounced in music, volitional control of structural characteristics (presumably interacting with awareness) become, once again, a prominent consideration[85].

Although this reasoning remains congruent with the most consistently reported results, we should also note some curious individual and task-dependent inconsistencies and some important details that remain unknown. Although often excluded from the main statistical analysis, some vocal participants (and one violinist [Hafke et al., 2016]) produce an initial following response rather than the more common opposing reaction. Larson and Robin (2016) argue that this relates to individual differences in motor control. A laryngeal EMG study, for example, revealed distinct and particular cricothyroid muscle contractions enabling opposing and following responses respectively (Liu et al., 2011). However, this does not necessarily explain *why* some individuals would operate differently. It is also not clear what the exact percentage of individuals is who preferentially display following responses. Not all pitch-shift paradigm studies report these details and those that do (see above; also Hain et al., 2001; Larson et al., 2007; 2008; Liu et al., 2009) report varying values between 2 and 30%. Presumably this relates to the differing methodologies used but this too remains to be explored further. One probable hypothesis is that the magnitude of the pitch-shift stimulus change correlates positively with the percentage of following responses (e.g. Burnett et al., 2007). Another curious finding that perhaps relates to this matter is that, although some form of initial response is automatic, we are able to follow instructions to alter voice F0 volitionally either in an opposing or following direction (Patel et al., 2014). Intriguingly, the following responses, given these particular instructions, had a much shorter latency (approximately 150 ms) than opposing responses (approximately 350 ms) with the latter exhibiting a small

---

[82] A recent study on violinists showed an almost complete compensation to perturbations in f0 auditory feedback of -48 to 48 cent. These corrections, notably, also occurred below perceptual thresholds (JND) of 7 cents (Hafke et al., 2016). It is worth noting from this study also that JND in violin playing was considerably lower than those reported in studies of singing (Grell et al., 2009; Jones et al., 2013; Hafke et al., 2013 - 10.5/15/25 cents respectively).

[83] Interestingly nearly half of amusics (4/9) compensated similarly to controls to perturbations of 25 and 200 cents (Hutchins & Peretz, 2013). Vocal pitch matching accuracy, but not the level of capability to perceive small pitch changes, correlated with the presence and size of the compensation to the perturbation. This further supports a dissociation between pitch perception and the production of pitch (see above) and that this initial correction mechanism occurs below conscious perceptual thresholds.

[84] Typically these studies are rarely explicitly directed towards musical production but it is, in my opinion, inherent in the task of sustaining a stable pitch.

[85] Although not reported in these terms another recent study also supports the idea that volitional control is related to the time-span of stable vocalisation/pitch. Keough and colleagues (2013) documented that trained singers could suppress compensatory responses to auditory feedback manipulation (between 2 cents – 100 cents) only if they occurred 500-2500ms after vocal onset. Earlier on (and throughout in non-singers) the adaptations were entirely automatic and beyond conscious control. This may perhaps reflect a process of discrimination of self from non-self-vocalization (see Behroozmand et al., 2011; 2016).

early following response[86]. Future research will hopefully shed some more light on these issues. However, it is also perhaps worth stating that the experimental paradigm, especially for shifts beyond normal motor variance, is somewhat artificial and thus may come with some inherent limitations in terms of its wider relevance to more naturally occurring pitch corrections in human interaction. Nevertheless, a music-specific pitch correction mechanism (analogous to period correction in pulse) is a distinct possibility (see below 'sensorimotor harmonisation' for further discussion).

Musical Pitch in Non-musical Human Interaction

In the following discussion I caution against interpreting studies of "pitch synchrony" and "tonal interaction" in spoken dialogue and mother-infant interactions as being a full analogy to pitch matching in music. Methodological and analytic concerns with the studies are highlighted, showing that the evidence is not as strong as proposed in the articles reviewed. Similarly to the above discussion on correction mechanisms, I argue that a new paradigm for investigating sensorimotor harmonization is needed to investigate these issues further.

We have explored comparatively the perception and production of pitch separately so far in this chapter. However it is the intricate interplay between the two in harmonious social interaction that is so clearly exemplified in music (e.g. Bannan, 2012). To talk about harmony is ultimately "to talk about the interpersonal processes of musical performance through which individuals volitionally and cooperatively blend musical lines in pitch-space and time" (Brown, 2007 p. 18). In music this may take individual and cultural form in, for example, sustained chanting, a persistent underlying drone, "chordal" homophony, or in a complex polyphony of individual "voices". Although these may appear very diverse I would argue that they share a common architectural foundation. I have suggested, just as capacities to engage with a basic musical pulse ground a wealth of complex rhythmic structure, that musical pitch in all forms is built upon a primary ability to produce and engage with a sustained stable fundamental frequency, and the ability to create or process certain relationships between pitches. This is built on the hypothesis that musical pitch structure worldwide — whether monotone or florid — can be characterized as being relationally organized with reference to sustained yet variable tonal areas (McAllester, 1971)[87]. This may appear controversial to some but, importantly, does not necessarily imply any particular form of designed tonal system or hierarchy. It merely suggests that musical pitch at any given time (or possibly within breath or phrase boundaries) is organized relationally within a framework of a dominant pitch region or regions (Bispham, 2009)[88]. Critically, in recent cross-cultural research, it has been shown that even in musical traditions featuring equitonic (equal-spaced) scales (e.g. East African Music) there is evidence to suggest that tonal centers are still perceived by idiomatic listeners (Ross & Knight, 2018). Broadly speaking the harmonic blending of voice or pitch in music (excepting music that is exclusively rhythmic) is clearly evident and is fundamentally integral to the structural composition of musical form. It is also this capacity and the structural management of pitch structure that affords harmonious and synchronised group interaction in music. A crucial question, therefore, is whether pitch intervals and 'pitch/tonal synchronies' are similarly operational in turn-taking and/or overlapping non-musical interaction.

"Pitch synchrony" in speech

Vocal imitation and repetition of prosodic contour appears to be functional in communication and vocal development from at least the third month of life onwards (Gratier & Debouche, 2011). Furthermore, emotional connotations of certain "musical" intervals in English and Dutch speech prosody have been argued to correlate to their most common association in modern-day Western speech. In particular major and minor thirds are more common in reading cheerful or sad passages respectively (Schreuder et al., 2005), and people re-enacting a sad utterance typically use a predominance of minor third intervals (Curtis & Bharucha, 2010). In spoken dialogue we can also observe interactional pitch structuring across turns. Boone (2003) first described "pitch synchrony" in an analysis of conversations between two chat show hosts. The pitch of their speech, when exchanging dialogue, was described to be often matched in unison or being consonantly related. A certain degree of functional intent in the quasi-harmonic structuring was suggested by the finding that an initially dissonantly related pitch was often chosen in interrupting and conveying disagreement. A follow-up study using human coding (artificially corroborated) of pitch intervals in pre-recorded television interviews found similar results as well as a predominance of unison pitch, perfect fifths, and octaves between the last utterance of one person's utterance and the first "note" of the conversational partner's response (Boone et al., 2007). Further and more focused evidence of interpersonal pitch structure correlating with socio-intentional function in speech has since been reported. Okado and colleagues (2012) presented results indicating that conversations that are primed to be agreeable exhibit more consonant intervals than those primed to be disagreeable[89]. Correspondingly in another study, comparing Spanish conversations between Chilean individuals given cause to be 'low-trust' versus 'high-trust', the clearest distinction was that the precision with which intervals matched "musical modes" was

---

[86] Another interesting finding that remains to be fully understood is that lengthy exposure to feedback that is slightly shifted over numerous trials can lead to more permanent adjustments even as auditory feedback is returned to normal (Jones and Munhall, 2000, 2005; Jones and Keough, 2008).

[87] Although I am not aware of any direct experimental evidence it is interesting to consider that, in fast-moving or florid passages, individual notes are likely to be too fast to be the subject for correction of discrepancies between vocal or instrumental production and an internal target or external reference (e.g. Lindblom and Sundberg, 2007). Thus, we can perhaps assume that these would be formed into a larger gestalt with the new whole constructed in reference to a more sustained and internally manifest pitch centre. This, in turn, perhaps suggests a more realistic target for adjustment.

[88] The human ability to recognise and manifest pitch centre in complex musical structures is a topic beyond the boundaries of this thesis but is discussed in Podlipniak (2016). He argues, with reference to the 'Baldwin effect' that this mental capacity emerged by joining the implicit recognitionn of the frequency of pitch occurrence, working memory and the emotional assessment of predicted stimuli.

[89] A particularly noteworthy finding here was that tritone relationships only occurred in the disagreeable condition.

higher in the high-trust condition (Robledo et al., 2016b). Finally an analysis of question and answer dyads found that the f0 in the final apex of the Question and the first of the Answer was more likely to correspond to Western musical interval categories when the Q+A pair's turn transition was periodic than when it was aperiodic[90] (Robledo et al., 2016a). Intriguingly, the data were determined to match more closely the equal tempered chromatic scale than other Western and non-Western tuning systems[91] thus intimating some cross-domain enculturation and learning. The authors concluded from this, in contrast to supposing temporal entrainment to "be a basic biological phenomenon essential to all types of cooperative joint action", that "mutual accommodation to pitch intervals is self-evidently learned by cultural transmission" (p. 1074). This conclusion, however, seems starkly at odds with the cross-cultural scope of the references cited above and related studies in the following section on tonal synchrony. A more likely premise is that the use of relative pitch cues and "pitch synchrony" form part a global capacity to structure socially intentional interaction with some culturally distinct specifics in terms of preferred intervals and tuning.

"Tonal synchrony" in parent-infant interaction

An interpersonal "tonal" pitch structure has also been proposed in parent-infant interactions both across turns and in overlapping vocalisations. Van Puyvelde and colleagues (2010) analysed musically transcribed interaction periods between Flemish-speaking Belgian mothers and their 3 month old infants. They proposed that, in brief sections of mutual vocalizing, both mother and infant adapted their utterances such that their pitches become tonally related or "synchronized". They further observed a dominance of consonant intervals that aligned with the harmonic or pentatonic series. In a later study, this time using acoustic analytic measures (Van Puyvelde et al., 2015), they found a similar proportional occurrence of tonal interaction periods (TIP) versus non-tonal synchrony (80-85%)[92] in Flemish and Mexican Spanish speakers. This suggested that this feature in phases of mother–infant vocal communication occurs across cultural and linguistic divides. Interestingly, in another related study, the authors also document significant relationships between tonal interaction periods and moments of social interaction repair and between non-tonal interaction periods and a loss of motivation and affect (Van Puyvelde et al., 2013). This, as they describe, highlights a potential importance of tonal synchrony in the attentional flow and regulatory success of these early social engagements[93]. As such they posit that it represents features of an ontogenetic correlate of an early phylogenetic "musilinguisitic" communicative stage in human evolution (prior to "prosodic protolanguage" [e.g. Brown 2001; 2017; Masataka, 2008; 2009]). They describe it as a border zone in vocal development "being neither musical nor linguistic but possibly offering a starting point to develop both musical and linguistic capacities" (Van Puyvelde et al., 2015 p. 51).

The measured scope of interpretation in the preceding quotation seems appropriate. It would, in my opinion, be premature to interpret the results of the previous sections on "pitch synchrony" and "tonal synchrony" as highlighting direct mechanistic analogies to pitch matching and harmonisation in music. The analysis is necessarily based, variously on normalised software data (e.g. Praat – e.g. Robledo et al., 2016), human analysis and interpretation (Boone et al., 2007; Van Puyvelde et al., 2010), and/or an analysis of peak and/or step accents[94] (see Liberman & Pierrehumbert, 1984; e.g. Van Puyvelde et al., 2015). These are all valid efforts given the constraints but do come with limitations in terms of a comparative discussion on music. Crucially, they all artificially "musicalize" and definitely "auto-tune" the data. Perhaps the most critical restriction, however, for our purposes of identifying specificities of music is the lack of an experimental reference point for how simple harmonic/tuning corrections are achieved in music. Do the findings reviewed in the previous two sections describe a looser manifestation of analogous processes found in music or do we use distinct correction mechanisms in musical interaction as was suggested by the auditory feedback studies (see above)? Contrary to suggestions made by the authors, the above discussion highlights that the current evidence for the former possibility is unconvincing.

Sensorimotor Harmonisation – A Missing Experimental Paradigm

In absolute terms we have impressive capabilities and potential to match perceived pitches. Sundberg (1987), for example, showed that trained singers can match a 440Hz reference tone with an accuracy of 1Hz (≈ 4 cents). Almost entirely not explored, however, are the processes with which we maintain and embody continuing pitch relations in music. Even in the simplest scenario of interactively matching a single sustained pitch we must necessarily be constantly correcting for efferent and afferent variances. One valuable exception to the general lack of experimental evidence is Grell et al. (2009). In this study choir singers were recruited to produce an F#4 (370 Hz) to complete a major chord with an external D4 (293.7 Hz) and A4 (440 Hz). Rising or descending shifts of quarter or semitones were introduced at varying points of a 5 sec sustained chord and were compensated for by the choristers. The authors reported that the singers responded in the range of 197– 259

---

[90] Following Ogden & Hawkins (2015) a segment was classified as periodic if the last 3 successive f0 pikes (transcribed from Praat tracker) differed by no more than ±15%.

[91] Crucially the small step size between intervals was analysed to not be the only influential factor in this assertion.

[92] TIPs where classified where vocal pitch accents of mother and infant are "related to their tonal center in the same way as the harmonics are related to their F0, without exception" (p. 47). In my opinion this is highly problematic. Their use of the harmonic series extends to H16 giving a total 8/12 "tonal" notes compared to the full chromatic scale for any chosen tonal centre. Given that the choice of tonal centre is additionally subject to interpretative best-fit I would suggest that this analysis needs to be confirmed in comparison to a randomised control.

[93] Curiously, and in stark contrast, results from a study on psychotherapy sessions indicated that higher levels of synchrony were related to poorer therapeutic relationships and greater distress (Reich et al., 2014). It is not clear why this was the case but perhaps a certain degree of dissonance / tension is reflective of a necessary degree of stimulation and interest (just as Robledo et al (2016) found that high levels of periodicity across question and answer turns was associated with low arousal).

[94] A 'peak' refers to a stressed syllable or an extreme edge of a variable contour, whereas 'step' would be classified as a longer plateau-like pitch period or contour with minimal fluctuation.

ms depending on direction and size of the pitch shifts, as well as on skill and anesthetization. However, observations from individual participants also indicated a fast but imprecise reflex reaction (presumably subliminal) at 50ms and some slow reactions, presumably under a degree of volitional control, at 723ms and 1290ms. These and many more details of how we achieve and maintain coordination of pitch frequencies remain to be confirmed, explored, and illuminated further.

I suggest that much could be realised from an analogous experimental design to the 'sensorimotor synchronisation' isochronous tapping task (Stevens, 1886) for rhythm and timing research. As we have seen in Chapter 2 (also Chapter 4) this seemingly simple paradigm has spawned a vast body of research that has informed our psychological and neuroscientific understanding of music, movement, action-perception coupling, attention, volition, interaction, sociality, empathy, co-operation, liminal and subliminal action, and group synchronization. I am, therefore, proposing some original pilot studies in which participants will be required to match a sustained external frequency (pitch) with a slider similar to that used in Hutchins and Peretz (2013; see above). Other production methods to be explored could include producing a tone by drawing a line on a simple touchscreen device, using a tensile string, and, of course, singing. The latter would naturally have the broadest interpretative scope and relevance but would also, given the degrees of variance and vibrato in singing, present the greater methodological and analytical challenge. Shifts in the range of 1-50 cents[95] to the target tone would be introduced and the first inquiry would be to identify thresholds for compensation and corresponding correction latency. Pitch processing thresholds are perceptually well understood (see above) but, in parallel with the tapping paradigm, it seems likely that embodied correction will occur even below liminal thresholds. On the basis of investigations with the voice feedback frequency shift paradigm on auditory control of voice F0 we can make some further initial suppositions and hypotheses. For example, it seems plausible that involuntary responses to perturbations will occur similarly at a short latency of circa 100-150 milliseconds and at a longer latency of circa 250-600 milliseconds with the latter being imbued with awareness and subject to attentional manipulations. If this were to be the case, we could perhaps surmise that the latter, being necessarily based on a more sustained pitch target, reflects a specifically musical pitch correction process. It is probably wise to resist temptation to think too far ahead. However, I would like to stress that, assuming initial success in the pilot studies, this paradigm would be open to extensions investigating, for example, different frequency relations and thresholds in the level of the shift at which corrections incur conscious awareness and/or are subject to manipulations of attention and volitional control. It would also be interesting to explore the human-human interactional dynamics of the task, and associations of this ostensibly simple task to issues such as empathy, prosociality, and cooperation, and the neuroscientific correlates.

Intervals, Scales, and Tonalities

In accordance with the scope of enquiry in this thesis and the intention to target mechanisms and capabilities that comprehensively root musical engagement we have, as in chapter 2, focused on a seemingly most architecturally foundational sustained tone. I would maintain that this is necessary in order to encapsulate all instances of musical pitch (including, for example, monotone chanting) whilst addressing the issue in terms of musicality representing an emergent capacity for musicality rather than a particular cultural and learnt form. However, certain higher-level universals in terms of the organization of pitch in music have been proposed in the literature (e.g. Justus & Husler, 2005; McDermott & Hauser, 2005) These have included the presence of unequally stepped scales (Sachs, 1943; Burns & Ward, 1999); a dominance of and "preference" for small-integer value intervals (Dowling & Harwood, 1986); and tonal hierarchies (Castellano et al., 1984; Krumhansl, 1990). I agree that these features would appear to be at least near-universals. However, they are not an absolute requisite for pitch-based interaction to be "musical" and can also all be understood to reflect variously psychoacoustic parsimony, inherent processing bias, processes of enculturation, and cultural construct. Studies on infant predisposition and musical development, in particular, support this assertion by indicating innate or early predispositions and constraints that have, variously, been argued to lie at the root of cross-cultural similarities across musical genres (e.g. McDermott & Hauser, 2005).

A relative preponderance of low-integer intervals in music across cultures is hardly surprising given that infants and adults find these easier to process (e.g. Schellenberg & Trehub, 1996; see above). It is, however, certainly neither prescriptive nor definitional of music's scope. The same can be said for Gill and Purves's position (2009) that component intervals of the most widely used scales throughout history and across cultures are those with the greatest overall spectral similarity to the harmonic series, and Sethares' hypothesis (2005) that tuning systems and scales co-evolved with spectral characteristics of a culture's dominant instruments. Scales are essentially a mental representation of regularity in pitch. The abundant use of scales that use two or more interval sizes[96] can, thus, be seen to afford an understanding of notes taking on distinct structural functions by relating differently to the other notes (e.g. Balzano, 1980; 1982). Within a statistical learning framework it is, as Thompson (2013) explains, "unnecessary to assume there is a specialized process in the brain that categorizes incoming tones as members or nonmembers of a scale (e.g. Krumhansl, 1985; 1990)" (p. 127). An interesting example of short-term implicit statistical learning is Loui et al. (2010) who showed rapid acquisition of grammatical structure using the artificial Bohlen-Pierce scale. Over time enculturation becomes more entrenched. Whereas infants are equally good at detecting changes in melodic patterns based on Western or Balinese scales, adults are much better with familiar scales (Lynch et al., 1990). Further evidence for a narrowing of processual flexibility comes from results indicating that infants are more readily able to encode musical scales accurately if the scale contains unequal interval steps (Trehub et al., 1999). By contrast adults

---

[95] It would also be interesting to explore wider intervals but I am suggesting this range initially to keep the task focused on tuning issues that may occur naturally in music performance rather than degrees of change that would more likely be processed as a discrete intervallic change.

[96] To be clear, although we can perhaps argue that all cultures use scales of this nature, not all scales used use unequal steps. The use of equal-tempered whole-tone and/or chromatic in some Western classical repertoire, for example, definitely contradicts any such assertion.

performed equally poorly, in comparison to accuracy ratings using a familiar scale, with both unfamiliar equal step and unequal step scales. Trehub and Hannon (2006, p. 81) summarise broadly that “although it is generally the case that long-term exposure to music leads to greater specificity in the registration of contour and interval information (Fujioka et al., 2004), musical knowledge occasionally interferes with the encoding of pitch relations, especially when the relations are unfamiliar or unconventional” (see also Schellenberg & Trehub, 1999).

Tonal structures and hierarchies are also learned through experience[97]. For example, in tonal-harmonic music psychological measures correlate strongly with the frequency rate and distribution of tones (see Krumhansl 2000). In two significant cross-cultural studies Western participants were able to extract some perceptual knowledge and experience of style-appropriate tonal hierarchies in North Indian music (Castellano et al., 1984) and Balinese music (Kessler et al., 1984). A more recent study also found that Western participants were able to identify distinctive features of two rāgas - Tori & Multani- in unfamiliar Indian music (Rohrmeier & Widdess, 2017). The most common interpretation is that listeners abstracted tonal hierarchies from these unfamiliar styles based on the distribution of tones in the music. This is supported by a study using novel composed sequences in which results were consistent with this view (Oram & Cuddy, 1995).

Just as with scales (see above) and metre (e.g. Trehub & Hannon, 2005a; 2005b; see chapter 2) comparing adult and infant capabilities related to melodic key membership clearly advocates enculturation to specific manifestations and a gradual narrowing of perceptual focus. For example, Western adults often fail to notice alterations to a melody if the changes are consistent with the Western key and harmonic implications of the melody. Infants, however, seem to be insensitive to key membership. They perform equally well if the changes are harmonically consistent or not and actually outperform adults on the harmonically consistent trials (Trainor & Trehub, 1992). In effect the adults' acquired and culturally specific knowledge of tonality (e.g. Cuddy & Badertscher, 1987; Krumhansl & Keil, 1982; Lamont & Cross, 1994; Trainor & Trehub, 1994) hinders their ability to detect changes that do not affect their learned means of understanding the inherent structure of the melody. In terms of sensitivity to more horizontal harmonic expectancies and structure, cultural learning can be evidenced as early as 1 year of age, if infants participate in parent-infant music classes (Gerry et al., 2012; Trainor et al., 2012), and develops gradually thoughout development[98]. The most critical point for our purposes is that these structural connections and expectancies seem to be predominantly learned and, hence, are unlikely candidates for being constituent parts of a narrow faculty of music across cultures[99].

---

[97] This does not, of course, necessarily imply that the capacity to learn scalar or harmonic structure is not in some form specialized for music. Consideration of this is beyond the scope of this thesis but the general opinion seems to be that this acquisition reflects broader capacities for implicit statistical learning (Tillmann & Bharucha, 2000; Rohrmeier & Rebuschat, 2012) and hierarchical syntactic structure processing (Koelsch et al., 2013) that are shared across music and language (see Rohrmeier et al., 2015).

[98] Principally, an understanding of chordal/harmonic progression appears to become accessible later in development and typically accumulates between about 4 and 8/11 years of age (Costa-Giorni, 2003). Even at 4 years old we preferentially choose harmonically appropriate accompanying chords to familiar tunes (Corrigall & Trainor, 2010) and show sensitivity in ERP brain responses to substantial (but not more subtle ones that adults differentially respond to) harmonic violations (Koelsch et al., 2006; Jentschke et al., 2008). The former finding may have been the result of familiarity rather than harmonic knowledge. However, 4-5 year old children certainly seem to have some perceptual access to relevant implicit knowledge. They are better able to detect changes in tonal vs atonal music (Trehub et al., 1986) and, like adults, more consistently detect those changes that go outside an established key structure (Corrigall & Trainor, 2010). Children between 5 and 7 years of age can recognize notes that violate implied harmony (Trainor & Trehub, 1994), and by 8 years of age show more subtle differential responses to different within-key notes that either contrast or support implied harmony (Cuddy & Badertscher 1987) . French and Australian children between 6 and 11 years of age have furthermore been documented to exhibit faster responses to timbral judgments with concordant ‘expected’ chords in simple chord sequences (Corrigall & Trainor, 2010). This shows clearly that, even stripped down to a horizontal analytic outline, culturally appropriate insights of harmonic structure are implicitly active by this stage in development.

[99] For discussion on the capacity to learn implicitly musical grammars and to what extent this is shared with language see Rohrmeier & Rebuschat (2012).

Chapter Conclusion – A Musical Tone

The psychophysiological and comparative discourse in this chapter has focused on a most plain concept of sustained pitch – a musical tone. It is worth reiterating that this is by no means intended to present a reduced view of music's many forms and practices. The use of pitch in music, across composition, culture and time, is immeasurably variable and fascinating. Beyond the constraints of our current approach, it is ultimately a source of individual character, cultural identity, playful exploration, socio-intentional drive, creativity, emotional expression and regulation. I maintain, however, that the capacity to engage with a "simple" musical tone universally grounds the organizational use of pitch in music[100]. This applies, in principle, from the most unadorned use of a persistent musical drone, through to a private rendition of a folk melody, to instances of its fullest potential across cultures in impressive group displays of choral harmony and instrumental symphony.

The comparative analysis in this chapter has, I submit, delved deeper than previous efforts to describe specific features of pitch in music. Discrete tones have commonly been presented as one of the few distinguishing characteristics of music across cultures (Cross et al., 2013; Koelsch, 2012; cf .Bispham, 2009). More precisely and inclusively, however, these can be inferred as resulting from, and being an expected acoustic correlate of, more deeply embedded processes. A musical tone is inherently a sustained physical and attentional practice that is rooted in particular correction mechanisms and a degree of volitional control. Initial experimental results further suggest that it is inextricably imbued with an awareness of a pitch-structural framework. Essentially a musical tone is neither primarily defined by, nor primarily experienced as, absolute values or discrete scalar steps. It can be more or less flexibly instantiated in relation to internal or externally influenced expectancies. More critical to its distinct nature than any absolute acoustic attributes or analysis is that it is understood and psychologically maintained in reference to explicit or implied pitch references or tonal centres.

As such a stable pitch is in many ways similar to a sustained pulse. They both provide a mutually manifest focal point and framework for individual experience and social interaction. In fact, although the argument took a slightly different course, the comparative analysis of musical tone in this chapter highlights remarkable similarities to the conclusions of chapter 2 on musical pulse. In both cases relational processing is a key and pronounced (if not absolute) difference between humans and other species[101]. Furthermore as we approach specificities of music within human communicative repertoires, it is correction mechanisms that permit sustained utterance and attention, a high degree of volitional control, and an awareness of architectural framework that appear distinct to the context of music, whilst also being universal foundations of a capacity for music across cultures. Of further critical importance is that, in both cases, it is precisely those features that distinguish the particular nature of individual experience with music that also afford group synchronous and harmonious interaction. Effectively configurations of musical pulse and musical tone provide an attentional structure for managing personal experience in an extended perceptual present - a continual phenomenological linking of the immediate past, the current moment ("now"), and future expectation – and a specific architectural framework for interpersonal communication.

A musical tone refers to sustained engagement with reference to implicit or explicit stable F0 pitch areas within a phrase, requiring conscious access to fine-grained pitch perception. Pitch, loudness, and timbre are independently variable in perception and production. Furthermore, it involves intentionally influenced frequency correction mechanisms based upon a goal to maintain certain frequency relationships with the output of self and/or others.

---

[100] To clarify this should not be understood to imply that all use of pitch in music is necessarily structurally significant. Nor should it be interpreted to suggest that, that in certain instances the "musicness" of a performance cannot reside entirely in the rhythmic pulse. As highlighted above (Section II – Introduction) it is my assertion in this thesis that music can be universally identified by engagement with *configurations* of musical pulse and musical tone. One or the other could be entirely absent, but not both.

[101] Most likely this reflects and follows a wider move towards increased sociality in humans and a motivation to understand others as ourselves – in terms empathetically of our own psychological and physical experience.

Chapter 4 - Musical Motivation

This chapter discusses uses and functions that are consistently the main motivational forces for engaging with music across cultures. These include emotional experience, expression, and regulation; intersubjectivity; social alignment; cultural belonging, ritual action, and communitas. All of these are essential to a *full* definition of music. However, these intrinsic motivations and their dynamic correlates in sound and action are not specifically enacted in music. As will be extensively discussed in this chapter, they emerge early in development and are ubiquitous features of human communication and culture. Therefore, in the effort to describe motivational forces that are unique to music, I argue that musical motivation is distinct in the inspiration to give these central human drives a particular and extended attentional structure within the confines of configurations of musical pulse and tone. Functionally, this provides extended phenomenological space, stability, and a degree of abstraction, intensity, and meaning, as well as providing a particular framework for interpersonal and group interaction and for ritual action. The chapter starts with a brief description of motivation and then describes the crucial distinction between intrinsic and extrinsic motivation. The latter, almost by definition, is not specific to any given behaviour and hence the argument in this chapter will be focused on sources of intrinsic impetus.

Motivation

Motivation literally denotes a cause to move and can be defined as “a modulating and coordinating influence on the direction, vigor, and composition of behaviour” (Shizgal, in Wilson & Keil, 2001, p.566,). It arises from a wide variety of internal, environmental, and social sources and is manifested at many levels of behavioural and neural organisation. Broadly speaking, motivation is the essential catalyst to a full spectrum of human and non-human action ranging all the way from quasi cause and effect hormonal influences on mating patterns in female rats (e.g. Bindra, 1969; McClintock, 1984) to higher levels of behavioural and neural organisation where motivational states interact with cognitive processes in influencing behaviour and attentional foci (e.g. Simon, 1993; Shizgal, 1996). It is inextricably linked with our emotional experience. Motivational states both influence emotional appraisals (e.g. Scherer, 2000) and can emerge from emotional reactions to stimuli in the form of modes of “action-readiness” – a physiological preparedness for context-relevant action (Frijda, 1986; Frijda & Zeelenberg, 2001). Some neurophysiological researchers have even suggested that the terms emotion and motivation reflect different positions on an intensity-graded spectrum of affect (Iverson & Fray, 1982, Simonov, 2013). Furthermore the term encapsulates root influences on volitional/explicit (e.g. Pally & Olds, 2000), non-volitional/implicit, goal-directed (e.g. Wiener, 1948; Ashby, 1961) and intentional action (e.g. Bratman, 1990; 1999).

Music and motivation

Despite the potentially wide-ranging theoretical implications, generic discussions on universal motivating factors for music appear to be largely absent in musical scholarship. Only in neurophysiological literature on music therapy do we find explicit reference to the need for a theoretical model of musical motivation (Unkefer & Thaut, 2005). A conceptual model of the inspirational forces that prompt and sustain engagement with music is an essential step towards a complete understanding of the affective nature of music (see, Thaut, 2002). Even more significantly for the purposes of this thesis, comparative perspectives are necessarily incomplete without consideration of motivation. As previously mentioned, some researchers (e.g. Trehub & Hannon, 2006; see chapter 1; cf. chapters 2, 3) have argued that infants possess the necessary capabilities to engage with music and that mature musical engagement results from an additional (albeit unspecified), ontogentically acquired motivation. Additionally, even if, *contra* to the conclusions of the previous chapters, non-human primates shared the necessary skills to behave or interact “musically” they may lack any instinctive reason to do so. This issue should not be accorded secondary importance in anthropological/primatological comparative debates. Arguments and/or empirical evidence that non-human primate species are able to perform a certain task *if* artificially motivated are often presented with the implicit implication that the behaviour in question is thus less likely to have played a uniquely significant role in the course of human evolution. However, the functional/adaptive significance of any given set of behaviours is as much dependent on the biological and/or cultural forces that motivate a species to perform and/or sustain them in appropriate contexts as they are on the capabilities that underlie them. This last statement clearly applies to musical engagement as much as it does to anything else. Motivational factors are necessarily essential components of a generic and universal capacity for music. In fact, it is noteworthy, that even among “amusic” individuals, who variably have difficulty processing the subtleties of musical pitch or in finding a beat (see chapters 2 & 3), a complete musical anhedonia remains remarkably scarce (Hirel et al., 2014; Mas-Herrero, 2014).

Intrinsic motivation and “play”

The psychological literature on motivation identifies two main forms of motivation – intrinsic and extrinsic. Intrinsic motivation is described in Ryan and Deci (2000, p. 70) as “the natural inclination toward assimilation, mastery, spontaneous interest, and exploration that is so essential to cognitive and social development and that represents a principal source of enjoyment and vitality throughout life” (see also Csikszentmihalyi & Rathunde, 1993). It is clearly evident in healthy children’s inquisitive, curious, and playful behaviours in the absence of rewards (Harter, 1981), and is most evident when infants are securely attached to a parent (Bowlby, 1979). Intrinsic motivation is according to Ryan and Deci (2000, p. 71) “increasingly curtailed by social pressures to do activities that are not interesting and to assume a variety of new responsibilities” (see also Ryan & LaGuardia, 2000). The term extrinsic motivation, in contrast, refers to the performance of an activity in order to attain some separable outcome which in turn can be characterised as a striving for reward (approach) or motivation to evade punishment (avoidance) (Norman & Shallice, 1986; Elliot & Covington, 2001). Music making can, of course, at times be extrinsically motivated by conformity, pride, privilege and other influences. However, the vast majority of musical engagement worldwide is intrinsically motivated – it is most commonly a positive and approached action. People will consistently report that their most enjoyable and meaningful musical moments have come when they felt they could 'just let go' while negative valence in music (and logically a subsequent curtailment of creativity and motivation) is nearly always the result of extrinsic pressures (Steptoe, 2001). Ultimately, of course, "we wouldn't do it if it wasn't fun" (Bispham, I. personal communication). In our efforts to describe specific features of music we must therefore, almost by definition, focus the search on sources of intrinsic impetus.

Use and Functions

In a sense, uses and functions, are an explanatory level removed from the concept of motivation. We can be intrinsically motivated without necessarily or consciously being driven by a functional goal. However, motivation is also unlikely to be disconnected from use and/or function. The following section, therefore, briefly discusses broad cross-cultural perspectives on uses and functions for music. I posit that the critical challenge for the purposes of this thesis is not to encapsulate the huge multitude of functions of music, but rather to address what motivates us to put *any* given character or social action into musical form.

Perspectives from ethnomusicology

A species' motivational drives need not necessarily be directly reflected in a behavior's observed physical or social value. Nevertheless, they are also highly unlikely to be entirely disconnected and it is most often parsimonious to describe motivation and function as different levels of analysis and explanation[102]. Merriam (1964), following Nadel and Baker (1930), distinguishes between 'use' and 'function' for music. The former, referring to the situations in which music is employed in human action and society, are inestimably diverse. Functions, in contrast, are described as concerning "the reasons for its employment and particularly the broader purpose which it serves" (p.210). These he lists as: 1. Emotional expression; 2. Aesthetic enjoyment; 3. Entertainment; 4. Communication; 5. Symbolic representation; 6. Physical response; 7. Enforcing conformity to social norms; 8. Validation of social institutions and religious rituals; 9. Contribution to the continuity and stability of culture; and 10. Contribution to the integrity of society. Nettl (1983) suggests that points 5 & 7-10 could perhaps be conjoined into "the statement that music functions as the symbolic expression of the main values, patterns, or themes of a culture (1983, p.15). Clayton (2016) consequently offers a more succinct categorization into the following supposedly irreducible functions: 1. Regulation of an individual's emotional, cognitive or physiological state; 2. Mediation between self and other; 3. Symbolic representation; and 4. Coordination of Action[103].

Addressing the 'top of the pyramid'

Nettl (2015) argues, however, that the distinction between use and function can at times be rather nebulous and proposes framing discussions, instead, as a metaphorical pyramid for imagining a society's world of musical culture[104]. In this imagined structure, uses can be seen to form the base of the pyramid, while proposed lists of functions fill the centre. Towards the upper end we can position statements regarding a principal function of music for any one society. For example, McAllester (1954), commenting on Navajo culture, states that "many of the usual functions of music…. are subordinated to an all-important function of supernatural control" (p.88). Another example is Merriam's (1967) conclusion that for the Flathead Indians music functions "as a means for expressing the fact that they remain Flathead no matter what changes in their life have occurred" (p.158). The very top of the pyramid - perhaps an unobtainable theoretical ideal - is "a single overriding function of music for all humanity under which all others are subsumed" (Nettl, p. 264). Attempts to term a principal function for music across cultures have notably included Blacking's proposal that the patterns of human interaction in music relate to patterns of human organization (1973); Lomax's resolve that the most favoured songs are operational in reflecting and reinforcing cultural norms of behaviour (1968); and Small's view that music "defends the central expressive interests of a society" (1987, p. 19). Nettl's own belief is that "what music ultimately does for humans, is to mediate between various forces: to control humanity's relationship to the supernatural, mediating between humans and other beings; to support the integrity of individual social groups and mediate between them; and to mediate in ways not possible with speech between individual and of individuals (2015, p. 267). It does this, he suggests, by expressing the relevant central values of culture in abstracted form.

Nettl admits, however, to being rather reluctant to suggest a single function for music. Indeed, none of the above suggestions relate to *all* music or all it relates to. They are also perhaps rather nebulous in character and scientifically speaking hard to subject to analysis and/or falsification. It may well be, as Clayton (2016, p.56) suggests, that "music's efficacy depends precisely on the indeterminacy of a single, "true" or underlying function". It is furthermore not clear what in the listed statements is thought to be specific to music. Going back to Merriam's list of functions, for example, none are necessarily specific to music and could equally apply to other arts, ritualistic activity or oratory. A lack of musical specificity in describing the interminably vast multiplicity of function within and across cultures is similarly problematic in evolutionary debate and is reflected in the many and varied adaptive and exaptive rationales that have been proposed. Tomlinson (2015), for example, argues that the central problem in music-as-adaptation perspectives is that they tend to

---

[102] Similarly a common function in any given culture or age may not necessarily reflect an ultimate evolutionary consequence or pressures.

[103] Another proposed list of the social functions of music is Koelsch's proposed 7Cs (contact, cognition, co-pathy, communication, coordination, cooperation, and cohesion) (see 2014).

[104] He also proposes a double-sided coin metaphor to caution interpreting appropriately about a culture's and analyst's statements on this topic (Ibid, p. 265-67).

assume a "unilateral explanation for a manifold phenomenon" (p. 33). Music, as described by Van der Schyff and Schiavio (2017), "takes on so many forms, involves such a wide range of behavior, and serves so many functions, it seems difficult to specify a single selective environment for it."

However, how about we try to approach the top of the pyramid from a different angle? It seems clear to me that music can effectively encapsulate and/or abstract all of our most central vitalities, social drives, and cultural reflections (in varying combinations). It can be utterly sublime but also, at times, entirely benign. The central question to my mind, therefore, is what motivates us to put any given character or social action into musical form? Does music really have any absolute and distinct connections to emotion, intersubjectivity, and/or cultural coordination? Why structure the "communicative musicalities" of our dynamic mediations with others and the flow of our individual experience around the specificities of a musical pulse and/or musical tone as described in the previous two chapters?

Music and Emotion: A special relationship?

The following sections on music and emotion focus on the available experimental evidence and theoretical models in the literature The main issues discussed are how best to characterise the complex relationship between music and emotion and whether there is anything specific about the ways in we express and/or regulate emotional states through music. Broadly speaking it seems that this is not the case. I posit, however, that the constantly renewing and sustained nature of attention to emotionally charged action and sound afforded by music makes it a particularly effective means of influencing and stabilizing mood. In accordance with this idea I suggest that the particular efficacies of music in terms of emotion can be directly related to the extended attentional focus afforded by engagement with configurations of musical pulse and tone.

> *Music moves us because we hear human intentions, thoughts and feelings moving in it, and because we appreciate their urgency and harmony. It excites motives and thoughts that animate our conscious acting and appraising of reality. It appeals to emotions that measure the effort and satisfactions, advantages and dangers of moving in intricate repetitive ways. Evidently a feeling for music is part of the adaptations of the human species for acting in a human made world; part, too, of how cultural symbols and languages are fabricated and learned (Kühl 2007(8); Gratier & Trevarthen 2008).* (Trevarthen, et al., 2011)

A special function for music is most commonly attributed to its association with "emotion"[105]. This has variously been expressed in the literature in terms of music being the 'language of emotions' (e.g. Cooke, 1959) or, somewhat similarly, in terms of music being particularly valuable in expressing emotions (e.g. Juslin, 2001). Interest in these connections has yielded a vast body of interdisciplinary investigation and exhaustive reviews are available in the literature (e.g. Sloboda & Juslin 2011; Eerolo & Vuoskovski, 2013; Juslin, 2016; Swaminathan & Schellenberg, 2015). Research strands include experiments, qualitative interviews; brain imaging and EEG; Field studies; questionnaires; music therapy; and ethnographic research; with each having provided some central and consistent findings[106]. Very broadly, we do seem to have reached a consensus, contrary to earlier cognitivist opinions (e.g. Kivy, 1990), that "music involves true instances of emotion" (Sloboda & Juslin, 2001 p. 91). Emotions in music are not distinct in self-report studies, expressive behavior, or physiological reactions (see Scherer, 1993) from those experienced in everyday situations. The literature on emotion studies is, of course, extensive beyond the boundaries of the present discussion (see e.g. Davidson et al, 2003; Manstead et al, 2004; Frijda et al, 2000; and Scherer et al, 2001)[107]. Nevertheless we can briefly note that emotions appear consensually to be characterized in terms of some key features. These include that emotions are functional in guiding behaviour and providing contextually relevant goals (e.g. Keltner & Gross, 1999); have behavioural, physiological, and experiential components

---

[105] Strictly speaking 'affect' is considered to be a more precise umbrella term for emotion, mood, feeling, arousal, and vitality (Oatley & Jenkins, 1996) and generally refers to the valence of an experience. Generically speaking therefore, the term affect is most precise in terms of describing the "emotional experience" of musical engagement. However, the terms seem to used interchangeably with 'emotion' being the more common in music research. See below for further discussion.

[106] Juslin (2005, p. 99) summarises the central findings from these strands of research in a table as follows (added references are signified *):
Experiments: Music produces differentiated effects on self-report, physiology, and behaviour (Davis & Thaut, 1989; Kallinen, 2004; Krumhansl, 1997; Nyklicek et al., 1997; Panksepp, 1995; Pigniatello et al., 1989; Pike, 1972; Vaitl et al., 1993; Waterman, 1996; Vaestfjall, 2002).
Qualitative interviews: Listeners employ music to regulate, enhance, and change qualities and levels of emotion. They show considerable awareness about the music that need to hear in different situations to induce particular emotions (DeNora, 2001; Gomart & Hennion, 1999).
Brain imaging and EEG: Listeners' responses to music involve subcortical regions of the brain that are known from previous research to be involved in emotional reactions (Altenmueller et al., 2002; Blood & Zatorre, 2001;; Peretz, 2001; Schmidt & Trainor, 2001; Brown et al., 2004*; Gosselin, et al., 2005*; Koelsch, et al., 2005*; Menon & Levitin, 2005*).
Field studies: Music influences consumer behaviour, helping, and interpersonal attraction and conflict (Fried & Berkowitz, 1979; Honeycutt & Eidenmueller, 2001; May & Hamilton, 1980; North & Hargreaves, 1997; North et al., 2004; Bailey & Davidson, 2005*).
Questionnaires: Music serves various emotional functions in everyday life: to change moods; to release emotions; as a source of comfort; to match current mood; as a source of enjoyment; to relieve stress, etc. (Behne, 1997; Juslin & Laukka, 2004; Panzarella, 1980; Roe, 1985; Sloboda, 1991; Sloboda & O'Neill, 2001; Zillman & Gan, 1997; Wells, 1990). Music therapy: Music facilitates the expression, identification, and experience of emotions; improves the control of one's own emotional behaviour; helps to 'trigger' emotionally-laden memories of past events; and helps to diagnose patients' psychiatric conditions (Thaut, 1990).
Ethnographic research: Emotional response to music is a 'universal' phenomenon, typically involving the particular forms of engaging with music may differ from one culture to another (Becker, 2001; Blacking, 1973).

[107] For an even more wide-ranging comparative discussion of emotion and mood in nonhuman animals see Mendl and colleagues (2010) and Panksepp (2010) and for a recent review of the cross-species comparative neuroscience of primal affective experiences in humans and related animals see Panksepp (2011).

(Scherer, 2000); have proximal elicitors (Le Doux, 1996); involve "action tendencies" (Frijda & Zeelenberg, 2001)[108]; and involve mechanisms of cognitive appraisal that evaluate an organisms environment (e.g. Lazarus, 1982; Scherer. 1999). All of the above have variously been accredited to listeners' emotional responses to music. However, some are most likely of greater specific relevance to the full range of world musics than others. Appraisal processes, for example, would appear to be of relatively marginal significance to the intrinsic act of musical engagement. Sloboda and Juslin (2001) instead appear to support a primary relevance of action tendencies arguing that "a theory of emotional responses to music in terms of 'musical expectations' has clear explanatory value in relation to Frijda's (1986) notion of emotions as a function of monitoring match and mismatch" (p.92). The embodied nature of music performance and perception as well as the lack of a definite functional goal would also suggest that it aligns more productively with accounts of direct connections between action and emotion.

Expression and induction in experimental research

Without wishing to diminish the many forms, assumptions, and arguments that have been proposed in philosophical and psychological discourse on music and emotion (see Thompson and Quinto 2011), it seems fair to summarise that projects have focused on two broad foci of investigation: the expression and recognition of musical emotions; and the evocation and induction of emotions (e.g. Cochrane et al. 2013). These can be understood as the 'where' and 'how' problems – as investigating musical emotions as being externally or internally manifest respectively (Schubert, 2013). As noted in Schiavio et al (2016) "(t)his juxtaposition of 'external' and 'internal' points of view has resulted in several influential frameworks (Fabian et al. 2014; Schubert 2013); and has provided important insights across a range of musically-relevant domains such as music therapy (Baker et al. 2007) and music performance (Scherer and Zentner, 2001)" (p. 2). Both have also provided theoretical frameworks for generating testable hypotheses and empirical results. Research into the expression and decoding of emotion has most commonly been interpreted in terms of adaptations of Brunswick's (1956) quasi-information-transferral 'lens' model of communication (Juslin, 1995; 2000). This has allowed researchers to isolate features of compositional technique, performance, and acoustic structure and investigate how these correlate to perceptions of semantically discrete emotional categories and/or dimensions of affect (e.g. arousal and valence). In terms of the induction of emotion, the most comprehensive and influential framework for synthesizing theory and findings from various domains is BRECVEM (Juslin & Vastfjall, 2008; cf Scherer's CPM-based approach [Component Process Model]; see Scherer 2009). This designates seven mechanisms through which music might induce emotions: Brain stem reflex; rhythmic entrainment; evaluative conditioning; visual imagery; contagion; episodic memory; musical expectance. These are generally designated as being neurologically distinct and appear roughly in order of ontogenetic development and availability to consciousness (see Juslin & Sloboda, 2013).

Despite the practical value of these models we should, however, be cautious not to assume that methodological rational and constraints reflect the reality of experience, motivation, and/or function. Expanded Brunswik lens models fail to account for the embodied nature of emotional experience (see Maiese, 2015; 2017) and musical engagement and risk an abridged, verbally categorisable account of the complex phenomenology of musical affect (see Davies, 2010). In turn, the notion of music 'evoking' or 'inducing' emotion connotes a slightly ethnocentric notion of a listener passively receiving music as a physical product. "Music" appears to exist separately from any human action. As noted in Schiavio et al. (2016) "(m)usic, indeed, is not a sentient being, which makes it difficult to imagine how it could feel and express emotions at all" (p. 3). Thus, this notion also cannot capture the enacted/interactional experience of music or fully reflect an evolutionary/functional account of communication in music (e.g. Cross, 2009; see chapter 1) nor its extension from animal models of communication or regulatory function in, for example, parent-infant interactions. The implicit assumption is that music is not essential for human survival and well-being. Schiavio and colleagues surmise, following a clear and comprehensive critique of the internal and external research foci that "the main issue that emerges here is that these approaches have difficulty addressing the actual experience of music, which arguably involves more than response processes, internal processing or detached aesthetic appraisals. Put simply, these theories tend to suspend the actual living experience of music in order to explain it; and, in the process, reduce it to various categories and loci" (p. 4).

An enactive alternative

They advocate instead a view on music and emotion that builds on an enactive/dynamic systems approach to the study of the mind and cognition (e.g. Varela et al. 2017; see also Ramachandran 2011[109]). Musical emotions are thus assumed to be episodes of experience associated with the ongoing process of maintaining adaptive, self-sustaining, dynamical stability. Crucially this perspective views all musical action (including listening) as being motivated and goal-directed and hence as being essentially emotive/affective. It "highlights the primordial necessity of musicality for embodied, pre-linguistic and emotional-empathic forms of understanding, communication and social cognition, beginning with the primordial interactions between infants and primary caregivers (Cross 1999, 2001; Krueger 2013; van der Schyff 2013)" (Ibid, p. 5). As such it connects directly to broad perspectives on primary intersubjectivity in developing social bonds in infancy (e.g. Trevarthen, 2002) and extends from an understanding of the interactive processes of participatory sensemaking (De Jaegher and Di Paolo, 2007). This argument is much more conducive to a functional evolutionary account of musicality[110] and I definitely

---

[108] They alter the probability of subsequent behaviour and prepare an organism for relevant action accordingly (Frijda, 1986).

[109] The latter refers to a collaborative paper that cautions against a purely mechanistic and inductive explanation of emotion in neuroscientific research.

[110] In a more recent paper Schyff and Schiavio (2017), extend the relevance of this perspective to an evolutionary perspective on the origins of human musicality. They argue that understanding the enactive continuity between mind, culture, and biology provides a potential way

recommend it to be considered in full (see also Schyff & Schiavio, 2017b). That said, it is, of course, more directly challenging for empirical investigation and is perhaps best utilized in interpreting and translating existing research findings into a more accordant understanding of a musical environment interacting with the developmental histories of the participants involved and a more dynamic affective-motivational process.

Music as emotion regulation

Another way of understanding the functional connection of music with emotion that deserves a greater theoretical and empirical focus is as a means of regulation. The concept of regulation is widespread in the cognitive sciences, including the domains of motivation; attention; action; and emotion (see Banfield et al., 2004). It refers to conscious and unconscious processes of monitoring and corrective control through which functional systems – ranging from the molecular level through to social and cultural levels – are maintained in line with preferred standards (see Schore, 2003; Davidson, et al., 2002; Trevarthen, 2005). Regulation has both been studied with regard to intra-individual processes – termed 'self-regulation'(see Vohs & Baumeister, 2016) and inter-individual processes – termed co-regulation (Fogel, 1993) and mutual regulation (Sroufe, 2000). Given the general importance of regulation in cognitive and affective sciences and in particular given the rapidly growing field on emotion regulation (see Gross, 1998; Gross & Thompson, 2007) the relative lack of a consideration of the regulatory nature of musical engagement seems unfortunate and is, most likely, a considerable barrier to progress in understanding the affective nature of music (cf. see Saarikallio, 2016). A highly developed example of a musician's ability to self-regulate her affective state is reported by Damasio (2000, p.50) who recounts his meeting with the pianist Maria Joao Pires as follows:

> Some years ago, the brilliant pianist Maria Joao Pires told us the following story: When she plays, under the perfect control of her will, she can either reduce or allow the flow of emotion to her body. My wife, Hanna, and I thought it was a wonderfully romantic idea, but Maria Joao insisted that she could do it and we resisted believing it. Eventually, the stage for the empirical moment of truth was set in our laboratory. Maria Joao was wired to complicated psychophysiological equipment while she listened to short pieces of our selection in two conditions: emotion allowed, or emotion voluntarily inhibited. In the condition 'emotion allowed', her skin conductance record was full of peaks and valleys, linked intriguely, to varied passages in the pieces. Then in the condition of 'emotion reduced' the unbelievable did, in fact, happen. She could virtually flatten her skin-conductance graph at will and change her heart rate to boot. Behaviourally, she changed as well. The profile of background emotions was rearranged, and some of the specific emotive behaviours were eliminated, e.g., there was less movement of the head and facial musculature.

As suggested by Joel Swaine (personal communication, 2006) this example, albeit perhaps an exceptional one, is one of voluntary self-regulation that is cultivated by a variety of skilled practitioners and performers, such as musicians (Persson, 2000) and meditators (e.g. Tang et al., 2007) who develop skills that may be more broadly encapsulated as components of mental skills training (Connelly and Williamon, 2004). It can perhaps act as a paradigmatic example for how music and musical practice can be used positively in managing individually our personal affective experience. Effectively, whilst welcoming the research shift towards understanding music as a social, communicative and interactive act we need to also be mindful of the fact that music can be an intensely meaningful individual pursuit, practice and immersion. Contrary to recent suggestions (e.g. Morley, 2013; Trehub et al., 2015) I do not think, therefore, that we should accept the notion that *all* music (e.g. even private practice) is necessarily socially motivated. That said it is very notably the same attentional architecture that identifies the specificity of music in individual experience that also affords group participation, synchrony and harmony (see below). All musical behaviour relies on mechanisms that are inextricably embedded in systems designed for communication and the regulation of affect. The sense of sociality often felt in individual practice is, therefore, not necessarily an imagined social 'other' but can perhaps be described as a state of emotional and social openness and interactional potential.

Swaine has proposed, in efforts to develop a model of singing and emotion, that *"music regulates affect and expresses a strategy for the regulation of affect"* (CMS seminar presentation, 2006)[111]. Although a different formulation this shares the strengths of the enactive/dynamic systems approach described above. Regulation broadly describes the emotional effects of interactional and personal engagement with music without disconnecting between "listeners" and "performers"[112] – the "where" and "how". It also provides consistent and functional links to relevant literatures in other disciplines. Co-regulation of affect is, for example, paradigmatically exhibited in the communications that occur between patient and therapist in music therapeutical contexts (Ansdell & Pavlicivic, 2005). It is also widely assumed to be, throughout early development, the principal function underlying the interactions and protoconversations between infants and caregivers (e.g. Malloch, 1999) in which dynamic processes "balance changes in self-regulation against the need for collaborative regulations of relationships with other persons in various degrees of intimacy (Porges and Furman, 2011; Carter and Porges, 2013)" (Trevarthen and Delafield-Butt, 2013, p. 4).

---

through what they regard as a "problematic" dichotomy between adaptation and non-adaptation (see ibid).

[111] My understanding is that is that through the use of the term 'strategy' this phrase incorporates that notion that music is partially volitionally controlled (see chapters 2 and 3). The term connotes a degree of stabilisation and voluntary control.

[112] As mentioned previously this distinction is cross-culturally mostly invalid and anyhow highly artificial as "passive" listening is an active process.

Music and vocal communication: A common code?

The affective experience of engaging musically is, of course, highly complex and may contain some context-specific features or phenomenological states (see below). However, reduced to verbalisable categories of emotion or ratings of affect-relevant dimensions, music seems not to be overtly distinct from other forms of human (and to a large extent nonhuman) communicative repertoires. It shares a common code with the wider category of vocal communication. An influential and much-cited meta-analysis of 145 studies (Juslin & Laukka, 2003; also Thompson & Balkwill, 2006), confirms that decoding accuracy for broad emotion categories[113] is above chance in music and vocal communication, and suggests that vocal and music "expression" involve similar structural cues such as overall F0, pitch level, rate, and intensity. In other words it appears - in accordance with suggestions of a common biological grounding (e.g. Spencer, 1857; Scherer, 1995) and musi-language evolutionary hypotheses (see Chapter 1) - that both music and emotionally expressive speech employ aspects of a common motivational-structural code (Ibid). Unsurprisingly, for example, emotions characterized by high or low levels of arousal (e.g. anger/happiness or tenderness/sadness) typically use fast or slow rates in music and other forms of communication respectively (e.g., Gagnon & Peretz, 2003; Juslin & Lindström, 2010). Angry utterances are naturally delivered with high vocal intensity in speech and a correspondingly high dynamic in music. Another foreseeable recent comparative finding was that speech and music expressing happiness used larger pitch intervals than those expressing sadness in both South Indian and Western music (Bowling et al., 2012).

These basic, presumably biological and/or physically grounded, cues are available to young children and across cultures. They are operational in infant-directed speech and singing from infancy onwards (e.g. Trainor et al., 1997; Rock et al., 1999). 5 year-olds, for example, perceive and effectively decode tempo as a cue to emotion (Dalla Bella et al., 2001). Instructed to convey emotions through singing, 5- to 8-year olds instinctively modulate basic acoustic cues such as tempo, pitch, and loudness in their performances (Adachi & Trehub, 1998)[114]. It is presumably also these affect-structural correlations that allow listeners to identify emotions in music from other cultures with a reasonable level of success. Expressions of happiness, sadness, and anger in Hindustani (North Indian) pieces are correctly identified by non-enculturated Western listeners (Balkwill & Thompson, 1999) and by Japanese listeners (Balkwill et al., 2004). Conversely the same categories of emotion are identified above chance by the Mafa from Cameroon (with no previous exposure) in response to Western music (Fritz et al., 2009). Finally, Laukka and colleagues (2013) confirmed that Indian, Japanese, or Swedish listeners can identify emotions expressed in Western, Hindustani, Japanese, or Swedish music. In general, it seems that the expression of more 'basic' emotions is more likely to generalize across cultures. In the latter study (ibid.), for example, results are more congruent for basic emotions (e.g., anger, fear, happiness, and sadness) than for complex emotions such as solemnity and spirituality (Laukka, et al 2013)[115]. Of course, even in the perception of basic emotions we do not rely solely on biological cues. Listeners in all of the above studies performed more 'accurately' (i.e. their answers were more closely aligned to the compositional/performative intent) in identifying emotions in familiar styles of music. This is presumably a result of culturally learned musical expectancies being either fulfilled or violated (Huron, 2006) and is also reflected in the notion that music from our own culture will tend to be more emotionally rewarding that music from a foreign culture. For example, participants from rural India rate Western music as "tenser" than Indian music (Wong et al., 2009) while American participants conversely rate Indian music as being "tenser" than Western music (Wong, et al., 2011).

A recent direct comparison of acoustic cues in music and speech in terms of valence, energy arousal, and tension arousal[116] (Ilie & Thompson, 2006) supports the conclusions of the comparative meta-analysis (Juslin & Laukka, 2003) with regard to intensity cues in all dimensions and the effects of rate/tempo on subjective measures of energy-arousal. However, pitch shifts only affected tension-arousal ratings in musical cues and correlations between rate/tempo and measures of valence and tension-arousal differed between music and speech cues. Manipulation of rate influenced valence judgments for speech, but not music. The reverse was shown in terms of tension-arousal ratings. Correlations between these cues and tension-arousal ratings may reflect intrinsic (and culture-specific) features of musical structure and correspondingly emergent expectancies and feelings of tension and release (see below). Additionally, the evidence that valence ratings were not affected by tempo changes in music may be explicable in terms of them being less psychophysiologically indicative due to a greater volitional control and awareness of rate and tempo in music (see chapter 2). Alternately it may be that the arousal-correlations of tempo in music relate also to our internal matching of action-relevant dance patterns and differing hierarchical interpretations of the appropriate beat level (see London et al., 2016; 2017 CMS seminar presentation).

Initially more puzzling, and counterintuitive given the above discussions, however, are findings that pitch changes had opposite effects on valence for music and speech (increasing with pitch rises in speech and vice versa in music), and affected subjects' measurement of energy-arousal only in speech. That pitch-levels can embody a range of valence is perhaps not

---

[113] Unsurprisingly, both descriptive-categorical and bi-dimensional accounts of affect in terms of arousal and valence show considerable correlation (e.g. Russell, 1980). The question of which of these best describes the correlations between stimuli and reported affect remains a matter of considerable debate. The evidence would appear to support a primacy of dimensional correlates (e.g. Laukka, 2004; Laukka & Juslin, 2005). Following Faith and Thayer (2001) – who argue with reference to individual affective state rather than communicative inference – discrete emotion categories are perhaps best described dynamically as attractor fields within parameters of affect and valence.

[114] By contrast culture-specific acoustic cues to emotion such as changes in mode (major or minor) only become relevant after 6 years of age (Dalla Bella et al., 2001).

[115] In Balkwill & Thompson (1999) Western participants also, congruently, had difficulty recognising 'peacefulness' in North Indian music.

[116] The authors explicitly frame their experiments within Schimmack and Grob's (2000) three-dimensional model of affect that integrates Thayer's (1978, 1986) multidimensional model of activation with Russell's (1980) circumplex model and arousal and valence. As discussions on expectancy, "tension" and "release" in music (e.g. Meyer, 1956) are abundant (see below) this may well provide a better model of affect for studying musically significant affect. However, it is as yet not common-place in the literature.

surprising given that high or rising pitch, for example, can signal both deference (Bolinger, 1978; Morton, 1994; Ohala, 1994) and surprise, anger, and activity (Scherer & Oshinsky,1977) across taxa. However, this in itself does not account for the differential results that appear significantly to contradict a common motivational-structural code in music and vocal communication. Beyond positing the existence of domain-specific processes for decoding emotion, the authors do not account for these findings or the different effects on energy-arousal ratings. One possibility is that these differences reflect fundamental differences in what might be termed the "production code"[117] for music as opposed to speech and vocal communication more broadly. Unlike speech, and certainly unlike non-human animal vocalization, instrumental and/or vocal music necessarily involves a high degree of independent variability of frequency, timbre, and loudness (see chapter 3 for detail). Hence pitch in music may be less *independently* indicative of arousal than elsewhere[118].

The specificity of "musical emotions"

Despite a wealth of sentiment and anecdotal accounts that musical emotion is in some way unique or special, it has so far proved remarkably difficult to pinpoint an exact and specific nucleus or address why that may be the case. Sensations of "peak experience" (Maslow, 1968), "flow" (e.g. Csikszentmihalyi, 1990)[119], and "shivers" or "chills" (Panksepp, 1995) have all been reported as commonly felt descriptors for particularly strong emotional experiences with music (see Gabrielsson, 2001). However, these experiences, whilst relevant and suggestive, are perhaps not the best ground for a generic comparative perspective on musical emotions nor are they described exclusively in reference to musical absorption.

The notion that intrinsic features of musical structure, such as tonal systems (e.g. Bharucha, 1994), create expectancies based upon culturally constructed knowledge, statistical regularities, and "Gestalt laws" was first formulated in a theory of emotion in music by Meyer (1956) who related expectancies to feelings of "tension" and "release" (possibly analogous to the terms "mismatch" and "match"). Since then it has received theoretical and empirical support from Narmour's (1990) model of melodic expectations and Sloboda's (1991; 1992) experiments on the relationship between structural features of music such as enharmonic changes or syncopations and the sensation of "chills". Additional support has come from physiological and neurological studies into "expected" and "unexpected"[120] harmonic progressions showing increased electrodermal activity and heart rates (Steinbeis, et al., 2005) as well as activations of the orbital frontolateral cortex (OFLC)[121] in response to "unexpected" chords. Meyer's theory and subsequent developments of it (e.g. Huron, 2006) would certainly appear to account for important aspects of musics' relationship with emotion. Importantly the general priniciple of implicitly learned expectancies could be argued to generalise cross-culturally and relate to the specific quasi-architectural features of "musical pulse" and "musical tone" (see Chapters 2 and 3). However, short-term expectancies, for example, in the form of future-directed attentional pulses (Jones, 1976), pervade all manner of ecological and social engagement (see chapter 2) and may, thus, again not be able intrinsically to account for a putatively specific emotion-related functionality for music.

A further possibility is that a special relationship is inherent in the dynamic nature of music. One way of interpreting this is that musically relevant emotions may be described in terms of 'vitality affects' (Stern, 1985). Intuitively this approach seems relevant and certainly accords with evidence supporting common mechanisms in engagement with pulse in music and movement (see chapter 2). However, it remains an empirically somewhat indeterminate term and applies broadly to the whole of our affective experience and communicative interactions. A more empirically solid assertion is that musics' dynamisms allow emotions to unfold and intensify over time. A psychophysiological study (Krumhansl, 1997) investigating the time course of physiological responses (including cardiac, vascular, electrodermal, and respiratory functions) to emotionally-laden ("sad", "fearful", and "happy") musical stimuli showed significant correlations between the majority of recorded physiological responses and time. In support of the above hypothesis, the strongest physiological effects for each emotion generally tended to increase over time. Neurological fMRI studies (Blood & Zatorre, 2001; Koelsch, Fritz, et al., 2006) have also been thought to support the hypothesis by showing that the activity changes of neurophysiological correlates of emotion are increased in the second of two 30sec listening blocks as opposed to the first (see Koelsch, 2005).

Another, perhaps parsimonious way of making sense of the evidence is that sustained engagement with the emotional nature of music does not merely intensify emotional experience but is also operational in influencing *mood*. Music, in potential

---

[117] Gussenhoven (2002) identifies three universal and putatively biologically functional (i.e. motivational-structural) codes in speech: The frequency/size code which indicates anatomically relevant information regarding the signaller; the effort code relating to degrees of energy expended in the production of speech; and the production code that – borrowing from Liberman's concept of breath groups in speech (e.g. Liberman et al., 1970) - associates high pitch with the beginning of phrases the time. Correlations between emotional state and the frequency of vocalization thus results from different levels of energy generation during exhalation phases and most likely reflects a wider correlation between pitch and intensity in speech (and quite possibly a highly linear relationship in non-human primate vocalisation [e.g. Demolin, 2006]).

[118] This may also underlie another exceptional finding that, whereas negative emotions such as fear and anger are typically expressed by high pitch in speech (Juslin & Laukka, 2003), lower overall pitch supposedly makes music sound less "pleasant" (Jaquet, et al., 2014), and that low pitches are often used to to express emotions with negative valence in music.

[119] The experience of flow is, of course, commonly described as a highly meaningful and desirable result in music. It is certainly part of music's eudemonic association with physical and mental 'well-being' (see Huppert and Baylis, 2005). It is charactised as being a decidedly pleasurable state but it is not definitively distinguished in degrees of hedonic experience - rather by optimally sustained levels of concentration (Nakamura & Csikzentmihalyi, 2014). This could again be argued to relate directly to the sustained attention inherent in the specific features of musical pulse and musical tone identified in chapters 2 & 3.

[120] To be precise these should be understood in terms of chords that violate cultural constructed statistically relevant "rules" regarding the "correct" progression of harmony.

[121] Koelsch (2005) explains that the OFLC is a paralimbic structure that plays an important role in the processing of emotion (p.414). See Mega, Cummings, et al. (1997) for a neurological review of the limbic system.

contrast to other forms of adult vocal communication, is consistently and purposefully employed as a means of altering individual and group mood (Bailey & Davidson, 2005; DeNora, 2001; Gomart & Hennion, 1999). In fact, one could feasibly make the case that this is the most prevalent use of music in society (Ibid). Despite lacking a clear theoretical underpinning, music's ability to function at this level is so well acknowledged in most areas of psychology that experimental paradigms regularly use music as a means of inducing particular moods in participants based solely on the assumption and observation that it works (see Albersnagel, 1988; Gerards-Hesse et al., 1994). It is notable that although other forms of communication, of course, partially embody this regulatory function, music does appear to be particularly efficacious.

Unfortunately, however, mood is all too often conflated into a general concept of affect or simply left unspecified in research into music and emotion. Some influential researchers have even cautioned against a distinction between emotion and mood. Juslin and Sloboda, for example, argue that "(i)t remains unclear whether the distinction between mood (core affect [Russell, 2003]) and emotion, which is based more on folk theories than on scientific evidence (Beedie, et al., 2005), will survive closer scrutiny" (2013, p. 588)[122]. However, we can quite plausibly lay a claim of experimental neurorealism to this claim (see Racine et al, 2005; also Gruber, 2017) and, in my opinion, it is critical that we have a music psychological framework for distinctly discussing mood (even if the definitional boundary may necessarily be somewhat nebulous). Mood and emotion are widely differentiated by scientists and lay persons on the basis that moods last much longer; lack an identifiable or proximal eliciting stimulus event, are not necessarily accompanied by distinct facial expressions (e.g. Ekman & Davidson, 1994), and in terms of their consequences and functionality (see Ketal, 1975; Beedie, et al., 2005). With regard to the latter, Davidson (1994) coined the phrase "emotions bias action, whereas moods bias cognition" (p.54). This can be understood to mean that, whereas emotions arise in situations where adaptive action is required, moods denote shifts in baseline 'modes' of information processing – influencing memory, decision making, and evaluative judgements (see Bless and Forgas, 2000).

Sloboda and Juslin (2001) succinctly explain that "[t]hus emotions can be viewed as *phasic* perturbations that are superimposed on the *tonic*, affective background provided by the mood" (p.75, emphases in original). Emotions are not solely elicited on the basis of intrinsic features of an action or gesture but also upon the relationship of the signal to an individual's affective background - his or her mood (amongst other things). One possibility is that the regular and sustained attention to emotional action and sound afforded by music is particularly conducive to influencing, stabilizing, and/or resetting baseline affective states (i.e., Moods [see Davidson, 1994]). The most crucial point to highlight from the above discussion is that, fundamentally, music does not express, regulate, and/or enact emotion, in a way that is intrinsically unique. Broadly speaking, it seems that the particular efficacies of music in terms of our emotional experience can be directly related to the extended attentional focus afforded by engagement with configurations of musical pulse and tone. Although the scope of relevance will necessarily depend on the particular content, a sustained attentional framework enhances, regulates, stabilizes, and heightens meaning and memory. A similarly structured argument and conclusion can be made for music's effects on other cognitive capacities, social interactions, and intersubjective dynamics.

Music and Intersubjectivity: A special relationship?

Intersubjectivity through interaction is another critical motivating force for music. However, similarly to the previous section on music and emotion, I will argue that the way in which intersubjectivity emerges through the act of musical participation is not intrinsically specific to music. The use of movement, motive, and gestural form in forming attachments with others is shown to characterise early interactions in human development. Furthermore, the constituent parts of our generic capacity for culture, which undoubtedly are of paramount and integral significance in music, are operational much more broadly in human sociality (and to varying degrees in non-human primate interaction). They, also, are an embedded part of, but not unique to, music. If music appears to offer particularly potent intersubjective insight it is, I posit, a direct consequence of the sustained attentional framework afforded by the specific features of configurations of musical pulse and tone. The section immediately following discussion on the human capacity for culture supports this further by arguing that music, broadly speaking, does not provide an unusual degree of shared understanding between ensemble musicians. Finally, similar arguments will be made for the connections between synchronous action and social connection. These connections, too, are shown to be a critical component in early parent-infant communication and to feature extensively and functionally in a myriad of non-musical contexts. Once again I will argue that, if these connections are strengthened in music, it is because the nature of the attentional structures in music affords continually renewing and reinforcing moments of mutual emphasis and accord.

Motive in movement and gestural form

Human gestures function in the creation of a unique cultural world where motives are shared in narrations of movement expressive of inventions in thinking and acting – organized sequences of actions that convey original thoughts, intentions, imaginary or remembered experiences – that are made, not only with logical order to perform some action in the world, but with feeling for the drama of changing experience (Gratier & Trevarthen 2008). Gestural expressions, with their innate timing and combination in narrations, are the foundation for learning all the forms and values of the elaborate cultural rituals and the conventions of art and language (Gentilucci & Corballis 2006). Their motive processes expressed in sound are the primary matter of music (Trevarthen 1999; Gritten & King 2006). [Trevarthen et al., 2011 p.13)

---

[122] They follow this assertion with another stating that "(i)n any case, those affective changes that listeners identify with music appear to be intense and short-lived, occur during (rather than after) the musical event, and produce autonomic reactions" (Ibid, p. 588). As a generalizable statement I would argue that this is simply false and at the very least subject to the same charges of neurorealism as the prior claim.

As humans we are constantly motivated towards achieving intersubjective belonging and forms of attachment (Baumeister & Leary, 1995; Maslow, 1968/2013; Bowlby, 1969; 1973). Developmental studies on attachment and aspects of "musicality" clearly demonstrate these facets to be inextricably linked (Beebe et. al., 2005; Trevarthen, 2005). Intersubjective coordination is widely assumed to occur through multi-modally expressed gestures expressed in basic dimensions of time, form, and intensity (Trevarthen, 1993; 1998; Stern, 1985). Such "communicative musicalities" support all our interactions from infancy onwards. They manifest intentions in forms that can be sensed by other individuals as signs or symptoms of the motives that generate them (Trevarthen et al., 2011). According to embodied and enactive models of mind and theories they are additionally inseparable in perception and action (Varela et al., 2017). Thus every movement can be seen to carry communicative intent or potential. We need not posit the interactive "sharing" of affect as an additional stage to our feeling for the process and quality of movement nor the management and regulation of personal emotional experience. There is often no clear divide between actions intentioned to be communicative to an attentive other, and those that are entirely concerned with individual state and regulation[123]. Therefore, instead of considering problematic conscious nonconscious divides in infantile or ancestral gesture, Trevarthen and colleagues (2011) propose that we accept "gesture and its communication as inherent attributes of a fundamentally social animal endowed with a single will to move" (p.15). Mirroring suggestions for models of animal communication (Owings and Morton, 1998; Seyfarth & Cheney, 2003; see chapter 1) they argue further that "*all* communicative gestures of animals – head bobbing and hand waving of lizards, singing of whales or nightingales, cries of migrating geese, squeaking of mice, grunts of baboons – are both self-regulatory (felt within the body or guided by interested subjective attention to objects and events in the world) and adapted for social communication, intersubjectively (Darwin 1872; Porges 2003; Rodriguez & Palacios 2007; Panksepp & Trevarthen 2009They serve cooperative life, including adaptive engagements between species, be they predatory or mutually beneficial" (ibid. p. 13-14, emphasis added).

Forms of intersubjectivity in infancy

Albeit grounded in shared features of primate communication and understandings of animacy and intention we clearly have highly pronounced or species-specific motivations to share mental states, internal experience, and intentionalities – a capacity for culture (see below). These manifest and develop naturally in interactions with parents from birth and are indicative of a fundamental and shared ontogenetically emergent human capability that is independent of language. The idea of infant intersubjectivity is therefore "no less than a theory of how human minds, in human bodies, can recognize one another's impulses, intuitively, with or without cognitive or symbolic elaborations" (Trevarthen, 1998 p. 17). Ultimately models of intersubjectivity in infancy (e.g. Meltzoff, 2007; Stern 1985; Trevarthen & Aitken, 2001) extend their relevance into adult experience and thus go to the very heart of how we understand our feelings of self and other, human connection and relatedness, and human culture.

Stern (1985) argues against a notion of intersubjectivity being seen as a unitary construct and proposes instead that it be broken down functionally into interattentionality, interintentionality, and interaffectivity. Of these he views the latter – also termed affect attunement (Stern, 1993; 2010) - as being "the primary medium and the primary subject of communication" (Stern 1985, p. 133) during the first year of life. This is defined succinctly by Beebe and colleagues (2005) as "the cross-modal matching of intensity, timing and shape of behavior, based on dynamic micromomentary shifts over time, perceived as patterns of change that are similar in self and other" (p. 47). The infant is accordingly assumed to be able to appreciate crossmodal correspondences and to subsequently understand – implicitly and "automatically" - the quality of another's inner feeling state and to instinctively realize if this is shared. Following Bretherton and colleagues (Bretherton and Bates, 1979; Bretherton et al., 1981), Stern (1985) argues that at 9-12 months infants exhibit an axiomative change towards a discernment of having a mind of their own and that others have minds – quite literally a theory of mind. A further "quantum leap" (Ibid., p.124) towards mature intersubjectivity occurs as the infants realizes that his/her focus of attention, feeling state, and intention can be shared and that his/her own attentional focus can be similar or different to that of a communicative partner. At this stage "a mental self and a mental other can be interfaced, aligned, or misaligned. The partner's alignment is now a direct subject in its own right." (Beebe et al., 2005 p. 43) As such Stern's is essentially a view of how inner subjective states can be shared. "What is at stake…. is nothing less than the shape and… extent of the shareable inner universe (Stern, 1985 p. 151).

Trevarthen and Meltzoff in contrast focus initially more explicitly on the behavioural manifestation and translation of state and intention in motives of overt action which they view as being operational from newborns onwards. Meltzoff's theory concentrates on imitation and a perception of correspondence (Meltzoff and Moore, 1989) as a fundamental foundation for social cognition (Meltzoff, 2007). This he understands to be the primary source of understanding the movements of others as being "like me" (Meltzoff 1985, 1990; Meltzoff & Gopnik, 1993) – as a social mirror[124]. Trevarthen's perspective (1993; 1998) perhaps resonates most clearly with musical sensitivities in that intersubjective matching is understood to occur via innate sensitivities to the rhythmic expressions and the timings of movement forms. The intercoordination of inner

---

[123] Although the neurological processes of generating sympathy and imitative resonance are not yet fully understood, neuroscientific research indicates clearly that intersubjective communication of intention, attention and affect, are central components of our cerebral physiology and functioning, enacted through some form of mirror-neuronal resonance (Gallese 2003; 2007; Gallese et al., 2004; Braten, 2007).

[124] This notably accords with adult studies of imitation showing a correspondingly more interdependent or 'other-focused' information processing (Van Baaren et al., 2004) a heightened perception of closeness and an affordance of social connection (Ashton-James et al., 2007)

psychological states between individuals is conveyed and transmitted interactively via timing, intensity and spatial configurations. These are described as the "kinematic, energetic, and physiognomic parameters of others' movements… transmodal features of motivation, coded as emotions" (Trevarthen 1993, p. 123)[125]. Similarly to earlier discussion on emotion, the key point for the purposes of this thesis is that these connections between form, action, and intersubjectivity are not intrinisically unique to music. Music instead presents them in a particular attentional and interactional framework.

The human capacity for culture

Although we may of course debate the dividing lines and have varying interpretations of the evidence, it seems reasonably clear that at some point quite early in human development our motivation and irrepressible facility for intersubjective understanding and action go beyond those exhibited by other species. Our reality and assumed truths are markedly the result of shared cultural processes. Crucially, for the concerns of this thesis, these processes are an integral and necessary part of music but are, as will be evident in the following discussion, broadly operational in all human interaction. Tomasello and colleagues (2005) argued for a marked rather than qualitative phylogenetic shift emerging ontogenetically from dialogic interaction with caregivers at around 12-15 months of age in human infants[126]. In particular they isolate shared intentionality - the ability to participate in activities involving joint intentions and attention. Following comparative analysis they propose that "(h)uman children's skills of shared intentionality develop gradually during the first 14 months of life as two ontogenetic pathways intertwine: (1) the general ape line of understanding others as animate, goal-directed, and intentional agents; and (2) a species-unique motivation to share emotions, experience, and activities with other persons" (Ibid, p.675). The result, they argue, is a child's ability to construct dialogic cognitive representations and "participate in earnest in the collectivity that is human cognition" (ibid)[127].

Positing a complete lack of shared intentionality in non-human primates underestimates, in my opinion, the capacity of great apes (Boesch, 2005; Visco-Comandini et al., 2015; see below). However, apart from this, the evidence from experimental studies on infants and primates broadly supports their hypotheses. Tomasello and colleagues chose to structure the discussion according to gradually intensifying levels of understanding of intentional action. These can be understood as constituent parts of a 'theory of mind' (Baron-Cohen, 1997; Lombardo & Baron-Cohen, 2011). Whether or not a 'theory of mind' is shared with other primates depends, however, principally on where one chooses to set the bar (see Call and Tomasello, 1998) and therefore isolating individual components is indeed more productive for comparative analysis (see also [Livingstone and Thompson, 2009])[128]. As testified by the evidence reviewed in the previous footnote, our capacity for

---

[125] For a more comprehensive review and comparison of the models of infant intersubjectivity offered by Stern, Trevarthen, and Meltzoff see Beebe et al. (2005).

[126] They also argue for a generative developmental failure to acquire a full understanding of shared intentionality in autistic individuals. For a more fully developed perspective I would refer readers to a more recent paper by Trevarthen and Delafield-Butt (2013).

[127] They notably extend the relevance of this perspective to an intriguing and ambitious evolutionary hypothesis arguing that "the result of participating in these activities is species-unique forms of cultural cognition and evolution, enabling everything from the creation and use of linguistic symbols to the construction of social norms and individual beliefs to the establishment of social institutions" (Ibid., p 675). The peer-review section of the paper offers some lively discussion and varying opinions on this.

[128] Within a few months after birth infants appreciate biological motion (Bertenthal, 1996). They understand animate action and are widely interpreted to exhibit an understanding of human actions as being goal-directed by 6 months of age (e.g., see Woodward 1999). For example, infants of this age expect human hands, but not resembling inanimate objects, to reach for objects they were reaching for previously (Woodward, 1998). Nonhuman primates also appear to understand others as animate agents. They are, for example, able to predict the behaviour of others in familiar situations and attempt to influence their behavior through communicative cues (Tomasello & Call 1997). In the same experimental habituation paradigm used with human infants by Woodward (1998) some monkeys expect similarly for people to continue to reach for an object that they have previously gazed at (Santos and Hauser, 1999). Recent functional brain research offers further support showing that the primate neocortex is excited to anticipate the regulation of motor activities in accordance with their goals, and a multimodal perceptual monitoring of an internal simulation of the completion of the action (Fogassi et al. ,2005; Pezzulo et al., 2008; Pezzulo and Castelfranchi, 2009; Hesslow, 2012; Gallese and Rochat, 2013).

Great apes and human infants at 9-12 months of age (but not at 6 months) additionally recognize both trying and accidental action – the pursuit of goals (Tomasello et al. (2003) cf. Tomasello & Call 1997). Seminal habituation studies by Gergely and colleagues (Gergely et al. 1995; Csibra et al. 1999, 2002) showed that infants typically habituate to the inferred target endpoint of a moving dot rather than the varying movement trajectories presented in the presence and/or absence of a causational obstacle. In an interactive experimental setting, infants at 9-18 months (but again not at 6 months) exhibit an understanding of intention in others. They are, for instance, demonstrably less patient when an adult keeps a toy for himself than whilst making a well-intentioned (but unsuccessful) effort to give it over (Behne et al., 2005) and 14- to 18- month-old infants will chose more frequently to imitate purposeful but not accidental actions (Carpenter et al., 1998). In an almost direct comparison with the study by Behne and colleagues, chimpanzees gestured more and gave up efforts to receive food sooner when a human was unwilling to give food than when he was unable (Call et al., 2004).

At least by the age of six months human infants follow the gaze of others. At 12-months this extends to more challenging scenarios in which the focus of adult attention is not visible to the infant (Moll & Tomasello 2004), indicating an understanding that the adult can perceive and attend to something beyond the infant's own perceptual reality (Caron et al. 2002). Once again the same conclusion can be reached for great ape species. Chimpanzees also follow conspecific gaze direction to external targets (Okamoto et al. 2002; Tomasello et al. 1998), exhibit signs of surprise and monitoring in the absence of a relevant target stimulus (Call et al. 1998; Povinelli & Eddy 1996; Tomasello et al. 2001), and follow the gaze direction of humans to objects behind visual obstructions (Tomasello et al. 1999). Chimpanzees, but not monkey species (Hare et al., 2003) also realise the consequences of another individual's view on their subsequent behavior in competitive contexts – the nature of their goal-directed action (Hare et al., 2000; 2001).

Very soon after their first birthdays, infants begin to understand that, in pursuing a goal, an actor may consider a choice of plans and will choose particular action plans based on an evaluation of the situation and extenuating constraints. In a rational-imitation paradigm14-

shared intentionality and pervasive motivation to share psychological states with others are certainly pronounced and exceptional. However, Tomasello and colleagues go further than this in asserting that, despite dyadic interactions (Maestripieri & Call, 1996) and some maternal gazing and social smiling (e.g. Mizuno & Takashita 2002) there are *no* observations of conspecific engagement in triadic protoconversations between adult apes and infants involving a shared goal (Bard & Vauclair 1984; Tomonaga et al. 2004)[129]. However, there is at least some experimental evidence of action coordination in cooperative tasks in chimpanzees (Povinelli & O'Neill, 2000; Visco-Comandini et al., 2015). The latter citation is particularly impressive as the task to move a common cursor on a screen necessitated the animals to modulate their behavior during execution *and* during planning. An even more significant challenge to absolute nature of the assertion in Tomasello et al., (2005) however, is the view that non-human primates exhibit a considerable degree of interintentionality in naturalistic (perhaps context-specific) behaviours. Boesch (2005) for example, counters that the group hunting techniques of the Taï chimpanzees are commensurable with the description of shared goals and intentions proposed. The most critical point to note to a comparative perspective on music is that, although these components of a human capacity for culture are an integral part of music, they are not unique to it and are even, to varying degrees, shared with other primate species. They are crucial components of a broad faculty for music but not constituent parts of a narrow faculty for music.

Shared understanding in musical performance

Describing interaffectivity in terms of the form, timing and intensity of gesture in sound and action – the "communicative musicalities" of interaction – and understanding our capacity and motivation for shared goals and intentionality one could perhaps expect music to provide particular or particularly intense intersubjective insight. After all, from early development onwards music is predominantly a form of joint attention (e.g. Sebanz et al., 2006) and a shared intentional act[130]. It is a space in which affective experience, regulation and all forms of intersubjectivity can be given particular focus and extended temporal space. It certainly can feel through performance that we get to know our collaborating colleagues and ensemble especially well – that we connect through the act of music-making. However, a particular degree of accord and shared reflection is not necessarily imitated in verbal accounts of musical performance. Schober & Spiro (2014) showed discursive evidence from joint improvisations by professional jazz musicians that a fully shared understanding is not needed for joint improvisation. Even some experienced listeners with similar musical backgrounds could at times interpret the same improvisations very differently (Schober & Spiro, 2016). Performers were, however, more likely to endorse co-performer's accounts of their thoughts and actions than statements about the improvised compositional structure (Pras et al., 2017).

Of course any form of ensemble music-making involves an initial shared intention to perform and act with others – broadly speaking we share in the goal of musical performance. However, once engaged, to what extent is intersubjective awareness necessary (and/or desirable) for successful and appropriate group coordination? Keller et al., 2016 (p. 26) argue that in rehearsed "scored-based" music "alignment is facilitated by *representations of performance goals* that are shared across ensemble members" (emphasis added). Therefore their view is one of collaborative rehearsal that is "typically geared towards establishing a shared performance goal, i.e., a unified conception of the ideal integrated ensemble sound (Keller, 2008; Williamon and Davidson, 2002)" (ibid, p. 5). The idea of the projection of a shared and "idealized mental representation" (Keller, 2008) is, however, highly problematic. It most visibly fails to generalize to an account of musical improvisation which Keller and colleagues (2016) seem to understand as being distinct. They draw a line stating that "in freely improvised music, co-performers eschew fully preconceived goals in favour of transient shared goals that evolve spontaneously through mimicry and other interactive social processes during live performance" (ibid. p. 5). However, improvisation is not entirely spontaneous and neither should the performance of composed "pieces" be inflexible. Instead

---

month-old infants appreciate the rationality behind the choice of an actor's action in terms of the focal intention and desired result rather than the specific action plan (Gergely et al., 2002). Six domesticated chimpanzees were similarly trained to imitate a human actor's efforts to operate a light or sound switch. Like the 14 month-old human infants they ignored the awkward action plan chosen (e.g. using a foot or head instead of hands) if, and only if, there was an obvious physical reason/constraint (e.g. the actor's hands were full (Buttelmann et al., 2007; 2008). In a later study Buttelmann and colleagues (2012) also showed that, following an unexpected event, chimpanzees could interpret a contextual shift in inferred human goals, even those displaced in time.

Between 12 to 15 months human infants have been described as moving from a predominantly passive form of joint engagement to being engaged significantly in activities involving coordinated joint engagement (Bratman, 1989; 1992; Bakeman & Adamson, 1984). They participate in interactions structured around coordinated action roles for achieving shared goals (Gilbert 1989; Searle 1993; 1995; Tuomela 1995) – a shared intention. At this stage, infants (14 months old) will prompt adults to reengage in halted shared activities, or may even perform an unwilling adult's turn for them (Ross and Lollis, 1987). In effect, they functionally appreciate the 'other role' in interactions and the potential for role reversal (Ratner & Bruner, 1978; Carpenter et al., 2005). At 18-month-old infants will, notably, also consistently initiate more play with adults who imitate their actions than with those who interact without action imitation (Fawcett & Liszcowski, 2012).

Infants at this age seem to have an understanding of aspects of selective attention based on an others' prior experience (Tomasello & Haberl, 2003). It is also at around this age that infants first attempt to actively generate joint attention with others through gestures such as pointing (see Holl & Meltzoff, 2011). Tomasello and colleagues (2005) conclude therefore that "(b)y 12 to 14 months of age, then, the triadic interactions child and adult around external entities appear as more "coordinated joint engagement," (p.683) (see also Liszkowski et al.,, 2004; 2006; Kuhlmeier et al. 2003). Accordingly they posit that "(c)hildren are thus engaging not just in cultural learning, which depends on understanding others as intentional agents, but rather, by formulating joint goals and intentions, they are engaging in full-blooded cultural creation" . (Ibid p. 683)

[129] See Carpenter et al. (1995) for evidence of some triadic interaction, limited in the authors' opinions in comparison to 18 month old human infants, between chimpanzees and bonobos in interaction with an adult human and some objects.

[130] The early significance of human intention and/or alignment is supported by evidence showing that children at 2.5–4.5 years of age are more likely to spontaneously synchronize with a drumbeat produced by a real person than with one produced by a computer-controlled stick or with a recording (Kirschner & Tomasello, 2009).

they simply represent differing degrees of creative freedom rather than any form of absolute distinction. As such issues with the notion of a preconceived ideal apply to all music.

Countering perspectives advocated in other recent research papers (Keller, 2008; Seddon & Biasutti, 2009; Biasutti, 2013) Schiavio and Hoffding (2015) argue that musicians playing together need not necessarily have an awareness of their co-players' subjective states nor an awareness of a shared common goal in order to perform[131]. Neither do they view these as being sufficient to account for the kind of awareness required for joint musical attention[132]. They structure and assess their assertions around interviews with members of the Danish String quartet claiming that "a pre-reflective, embodied and interactive layer of awareness is necessary in accounting for joint musical performance, while the presupposition of more reflectively demanding 'goals', and 'attention' often is not" (ibid p. 17)[133]. One quotation from the interview with the cellist of the group stands out in particular:

> I think I have an idea of where I'm going and then I think … for each tone you play, you could say it such that time was a long, long line and you had a lot of points on the line … so for each point you advance, then … it gives … it will give the premise for where the next point would be because the tone itself in a sense defines the next tone and so on because otherwise the phrase becomes unnatural. And therefore you're really in the tone you're really on and then it kind of gives itself how the next tone will become … You need a smooth (development) and therefore each tone, each and every vibrato, each and every oscillation, bowing, phrasing, everything builds towards how it is going to become and it is impossible to predict how it will play out. It depends on what you laid as ground (Fredrik Sjölin, ibid p. 13).

This descriptive quotation is, in my opinion, especially insightful and relevant. It appears to describe the experience of engaging with music in direct accordance with the specificities of a basic musical pulse and tone described in chapters 2 and 3. We are not, at the time of musical immersion, concerned with a mental representation of others, the communication of particular emotional states, or the complete acoustic outcome – the final "portrait". Instead we are experiencing a moment – a psychological present that is constantly extended in relation to the timing, tuning, and dynamic form of the immediate past and corresponding future-directed expectations and the ongoing generation of action and attentional focus[134].

Synchrony, entrainment, mediated interaction, and prosociality

One prominent candidate for a special connection of music among communicative contexts to intersubjectivity and connectedness is synchronous and/or entrained action. Music and other forms of entrained action necessarily create recurring moments of synchrony and a relative proliferation of moments of shared emphasis and notice[135]. In a practical sense interpersonal synchrony can facilitate the coordination and collective performance of certain cooperative goals (see Allsop, 2016). Additionally interpersonal synchrony has intense psychosocial and implications – it both generates and reflects moments of interactional significance and accord. Broadly speaking, moving in synchrony with others seems to encourage prosocial behavior and a sense of interpersonal connection and agreement. For example, Gill and Borchers (2003) showed that moments of physical/gestural synchrony transpired at precisely the point in negotiating a joint architectural task when points of agreement were reached. These moments both signalled and implicitly sealed moments of consensus from which the negotiating partners could move forward. Baimel and colleagues (2015) congruently describe behavioural synchrony, like the human propensity for imitation, as being "part of a larger suite of processes that allow for effective interpersonal coordination between physical bodies and minds (Chartrand & Lakin, 2013)" (p. 4)[136]. For example, similarly to experimental settings with music (see below), mimicking behavior has been evidenced to increase affiliation and

---

[131] Keller (2008) argues that joint musical attention in partially predicated upon sensitivity to the subjective states of others and a shared sense of meaning. Sedden and Biasutti similarly argue that "to reach empathetic attunement, musicians *must* decentre and see things from other musicians' musical points of view" (2009, p. 120) (emphasis added) (quoted in Schiavio & Hoffding, 2015).

[132] An interesting suggestion made is that *explicit* communication and attention to co-players become more significant when the "normal mechanisms of co-performance are under pressure or have broken down" (Ibid p. 12). Following earlier research by one of the authors they posit an inverse relation between increased absorption and attention to one's co-players (but not, of course, vice versa) (Hoffding, 2014). In a quotation that will, I believe, resonate with many experienced musicians the violinist from their study describes that "(w)hen you play the best, you are *not* communicating" (emphasis added).

[133] In a related study on dance it was suggested that coordination in salsa dancing partnerships relies primarily on synchronisation with the music, rather than adaptation to each other (Koehne, et al., 2016). The partner's coordination therefore is not seen to be reliant on a continual mutual exchange of "information" or the monitoring of preconceived joint action plan. In effect the musical beat, is being posited as a quasi-intermediary and predictable framework (although, of course, in a natural setting this would also be co-maintained and co-generated by [and with] the performing musicians).

[134] It is worth noting that this view and the arguments in Schiavio & Hoffding (2015) concur with recent proposals for a greater focus of attention on the manner in which people align words and thoughts, bodily postures and movements in our broader understanding of social cognition. Social cognition, as is comprehensibly represented in a recent paper by Galetti and colleagues (2017), "concerns the graded and dynamic process of alignment of individual minds, *even in the absence of a shared goal*" (p. 253, emphasis added).

[135] Some of the research in this section discussed sustained forms of entrainment in non-musical contexts (e.g. rocking, walking). Typically these have not been analyzed in terms of phase and period correction mechanisms (see chapter 2). It is possible, given a general lack of awareness of "pulse", that continual period corrections are not similarly operational as in music. However, if it turns out not to be case I would perhaps need to adjust my perspective on the specificities of a musical pulse to further state a dynamic communicative context and the pulse acting as a reference point for variably complex rhythmic action and/or perception rather than being barely and functionally instantiated.

[136] It is interesting to note that that some aspects of the significance of interactive synchrony can also be found beyond our own species. Synchronous surfacing in male bottlenose dolphins has, for example, been associated with alliance membership especially during bouts of 'intense' social behaviour (e.g. mounting, displays, chasing) with female consorts (Connor et al., 2006)

induce subsequent prosocial behavior (van Baaren et al., 2003; 2004), suggesting a fundamental adaptive function in fostering relationships through the dynamic forms of interaction (Lakin & Chartrand, 2003). Thus if we are to posit any socially significant specificity for music as a form of continuously entrained action it is, yet again, likely to be rooted in the recurring, relatively predictable and stable nature of an underlying configurations of musical pulse and musical tone.

Synchrony in infancy

It has been proposed that synchrony extends an infant's fundamental biological rhythms to social relation and connection (Lester et al., 1985; Wolff, 1967)[137]. Feldman (2007), for example, assumes that in human infants "the capacity to engage in temporally-matched interactions is based on physiological mechanisms, in particular oscillator systems, such as the biological clock and cardiac pacemaker, and attachment-related hormones, such as oxytocin" (p. 4). She and other authors on synchrony agree that one of its most primary functions is in enabling infant self-regulation through co-regulatory interaction with caregivers (Feldman, 2003; Field, 1994; Fogel, 1993; Tronick, 1989). As such it is a crucial formative experience in the maturation of brain circuits that support social engagement between 3-6 months; it supports cognitive, social and emotional development and could perhaps, in absence or distortion, be indicative of forms of developmental psychopathology (Feldman, 2007; Johnson et al., 2005; Jaffe, et al., 2001). The mechanisms through which biological synchrony promotes infant development are not yet fully understood. However, Feldman suggests two principal possibilities: The first, which accords with similar findings in other mammalian species (Hofer, 1995) is that "physiological synchrony functions to entrain the infant's biological rhythms, thereby providing an external regulation to the consolidation of cyclic physiological processes occurring in milliseconds (e.g., hearth rhythms), days (e.g., circadian rhythms), or seasons (e.g., seasonal hormonal release)"[138] (Feldman, 2011 p. 575). The second possibility posed (Ibid; Feldman, 2007) is that synchrony helps to form and consolidate a unique bio-behavioral attachment - a specific physiological and behavioural fit between parent and child (Barrett & Fleming, 2011) - that affords an infant's physiological, social, and emotional growth (Carter et al., 2005).

The first of these hypotheses is supported by a body of research that has established links between interaction synchrony and degrees of heart rate variability (rhythm sinus arrhythmia [RSA]) in both mothers and infants[139]. Both exhibit higher levels of vagal-tone suppression (signifying a more adaptive autonomic response to stress) to a still-face following synchronous play (Moore & Calkins, 2004; Moore et al., 2009). Likewise, recovery from the still-face correlated with maternal and infant RSA (Ham & Tronick, 2006) as did degrees of interactional synchrony and mother's and infant's RSA (Feldman, et al., 2010). Feldman and colleagues (2011) extended these findings through an analysis of the *second-by-second* coordination of physiology and behavior between maternal and infant's autonomic response and relational behaviour during social contact. Their results supported a coordination of heart rhythms through episodes of interaction synchrony between mother and infant. Notably, albeit perhaps not entirely surprising given the importance of breath and vocalization in regulation of the autonomic nervous system (e.g. Porges 2003; 2007), it was moments of vocal synchrony (as opposed to shared gaze or gestural synchrony) that produced the closest physiological fit between the mother and infant.

Entrainment in social development

Even though the capacity to actively coordinate action in reference to a musical pulse appears later in development (see chapter 2), sustained periods of synchrony in the form of entrained actions appear to be socially relevant even in early infancy. For example, 12-month-old infants (but not 9-month-old infants) are more likely to exhibit a social preference for an animate toy whose chair is rocking synchronously with the infant's own chair than one that is being rocked asynchronously (Tunçgenç, et al., 2015). At 14 months infants demonstrate increased helpfulness following a physicalised experience of interpersonal synchrony in a musical context[140] (Cirelli et al., 2014a). In follow-up studies the infant's increased prosociality, following the synchronous condition, was shown to extend to other affiliated (but non-participating) adults (Cirelli et al., 2016), but not towards neutral strangers who sat passively reading a book (Cirelli et al., 2014b). Finally it was shown that the positive effect on infant helpfulness did not necessarily depend on the synchrony being musically enacted (Cirelli et al., 2017). Rather the difference between synchronising with music or simply being bounced in synchrony with the experimenter but listening to nonrhythmic natural sounds was that the infants were subsequently faster to help and less fussy. The music seemingly acted as a mood regulator or distractor to help keep infants happy and "allowed them to fully experience the effects of synchronous movement" (Ibid. p. 319).

By the age of 4 these social connections are actively extended into a motivation to move in synchrony with adults in joint rhythmic activities or in 4-6 year olds' peer play (Tuncgenc & Cohen, 2018). Children between 2.5 years and 4.5 years of age are able to interactively adjust a drumming tempo to a beat outside the range of their spontaneous motor tempo (in contrast to their capabilities in a simple sensorimotor synchronisation setting [McAuley et al., 2006; see chapter 2]). Significantly they synchronise with greater accuracy in a social condition (Kirschner and Tomasello, 2009). However, in

---

[137] We should note that "synchrony" in this context and in Feldman's work is more loosely interpreted than it's meaning in a musical setting. It refers more broadly to actions in a shared time but the implications are nevertheless highly significant in music.

[138] As noted further by the author (Ibid, p.575) the optimal organization of such biological rhythms in the first months of life has shown to predict the development of self-regulatory and interactive capacities across the first years of life (Feldman, 2006, 2009).

[139] Higher levels of heart rate variability (within certain limits) have been proposed to be indicative of a degree of flexibility in the autonomic nervous system and healthy regulatory functioning in mammals and are linked to fundamental features of social communication (e.g. Porges, 2003; 2007). These correlations therefore suggest a direct influence of synchrony on the most fundamental mammalian mechanisms of communication and affect regulation through changes in vagal tone (see chapter 1)

[140] In these studies infants were held and bounced in a baby carrier facing toward another person who either bounced in or out of synchrony.

this age range neither German nor Brazilian children exhibited an increase in prosocial tendencies following a simple isochronous rhythmic task with another human (versus drumming along an artificial playback beat) (Kirschner & Ilari, 2014)[141].

Entrainment and prosociality in adults

From the preceding discussion we can clearly see that the fundamental effects and social correlates of synchrony in interaction are not autonomously unique to entrained actions and music. Nevertheless the repeating pattern of synchronizing actions through entrainment appears to strengthen the psychosocial implications in adults in many ways. Succinctly summarised by Cirelli and colleagues (2017) "adults who walk, sing, or tap together are later more likely to be cooperative, helpful, and rate each other as likeable. When adults walk, row, tap, dance, or sing in synchrony with one another, affiliative behaviors such as trust, cooperation, and ratings of likeability are encouraged" (p. 319). In experimentally controlled studies with adults four key aspects of prosociality correlate with synchronous movements and vocalisations. They appear with reasonable dependency, despite some inconsistent results, to increase prosocial behaviors; enhance perceived social bonding; improve social cognition, and to a lesser extent/less consistently increase positive affect (see Mogan et al., 2017 for a recent meta-analysis).

Most commonly the prosocial effects of synchronisation are understood to occur, similarly to imitated movement (e.g. Decety & Sommerville, 2003), via a blurring of the boundaries between self and other (Hove, 2008). In a group context this blurring can be interpreted to evoke a sense of "oneness" with the group (Swann, et al., 2012). Another common explanation of the mechanisms involved is that we experience enhanced connection with others (Schmidt & Richardson, 2008) by establishing a form of attentional union – we give increased attention to the social cues and actions of interacting partners (e.g. Macrae et al. 2008). Extending from this idea Baimel and colleagues (2015) argue that synchrony fosters cooperation by exploiting our everyday social cognitive reasoning about other minds. They posit that "by directing attention to others and their mental states, while decreasing the perceived psychological distance between individuals, behavioral synchrony makes us better able to reason about other minds and thus coordinate and cooperate" (p. 870)[142].

People acting in synchrony (e.g. walking, cups game, tapping) subsequently cooperate more in group economic exercises with others, even in situations requiring personal sacrifice and independently of generated positive emotions (Wiltermuth & Heath, 2009; Valdesolo et al., 2010; cf. Schachner & Garvin, 2010). Joint tapping has also been shown to increase compassion and the subsequent tendency to be helpful to a "victim" in a social game situation (Valdesolo & DeSteno, 2011)[143] as well as to increase affiliation towards interactive partners (but not towards physically present others after simply synchronizing with a visual target on a computer) (Hove & Risen, 2009). These correlations appear to work the other way around too and can also be functionally observed by outsiders. For example, individuals going into a synchronisation task with a pro-social mindset (either by disposition of character of through priming manipulation) exhibited greater levels of interpersonal synchrony (Lumsden et al., 2012). The perception of social rapport from an external (visual or auditory) perspective also appears to correlate with the stability of interpersonal coordination (Miles et al. 2009). This mirrors broader findings that the observation of interpersonal coordination enhances the perception of commitment to joint action (Michael et al., 2016). Additionally, intentionally entraining individuals are boosted in self-esteem and will experience greater self-other overlap with their partner following periods of synchronous compared to asynchronous movement (Lumsden et al., 2014). Subsequent performance on a perceptual speed judgment task has also been observed in participants rocking in chairs

---

[141] It is interesting and pertinent to consider this and related studies together with my proposals for the root specificities of music in terms of broader discussions on the relationships between musical training and the development of cognitive and social capabilities. In contrast to the later study, Kirschner and Tomasello (2010) presented evidence that children who play together in a musical versus a non-musical way are later more helpful and cooperative with one another. There have also been some interesting projects in natural settings showing that long-term group music-making interventions promotes social development and sympathy (Rabinowitch et al., 2013), even among infants (Gerry et al., 2012). In the former study music interventions significantly outperformed drama and other control activities in terms of transferable empathic capacities.

Nevertheless in a review of the literature, Schellenberg (2013) suggests that "nonmusical associations with music lessons appear to be strictly cognitive" (p. 531). In some sense much of evidence can appear to support this position. In a study of 6-11 years parents' reports of social skills was not correlated with the duration of musical training received (Schellenberg, 2006). Experimental studies of 1 year and 3 years of musical training showed improvements, compared to control in social skills (Schellenberg, 2004) or self-esteem (Costa-Giorni, 2004) respectively. In a study of 7 and 8 year old children musically trained children performed better than untrained controls. However, the key finding was that the advantage disappeared when FSIQ was held constant (Schellenberg & Mankarious, 2012). Standard music classes taught in schools also show equivocal results (Rickard et al., 2012; 2013)

How then can we understand this apparent disparity between studies? My understanding is that the specificities of musical pulse and tone (amongst other features) sustain attention and thus afford an intensity and highly conducive and memorable learning environment. However the scope of transferable skills and propensities depends on the content of the musical structure and training. Notably most of the null findings involved music lessons taught individually. In direct contrast the nature of the task in Kirschner & Ilari (2014) was pared down to simple synchronous drumming and was far less socially involved than the musical task in the earlier study with a positive result (Kirschner & Tomasello, 2010). Finally a recent study provides a compelling contrast to the equivocal findings from classroom training showing that weekly group music classes *encouraging high levels of interpersonal synchrony*, showed larger boosts in self-reported sympathy and prosociality compared to controls (Schellenberg et al., 2015). I would argue that the nature of musical content determining its transferable relevance applies also more general cognitive abilities. This form of learning may occur by virtue of the degrees of volitional control and sustained attention in music via an overarching relevance to executive functioning (e.g. Moreno et al., 2011).

[142] A third perspective is that the effects of synchrony on coordination are supported by an affective channel. Although this receives less support from the experimental literature its relevance may increase substantially in naturalistic and ritualistic settings (see below).

[143] However, unfortunately synchrony does not necessarily promote morally positive action. For example, synchronous activity also boosts compliance with requests to aggress with cooperating partner (Wiltermuth, 2012).

in synchrony (Valdesolo et al., 2010), as has a notable shift from self-memory advantage (in anti-phase coordination) towards a balance of self- and other-memory when movements were coordinated in-phase (Miles et al., 2010). Finally, although not necessarily entrained, an increase in positive affect has been found to correlate with the physical and gestural synchronies in a study of cooperative, competitive, and 'fun' tasks through verbal interaction (Tschacher et al., 2014)[144].

Entrainment and group affiliation

Crucially these various effects seem to translate and generalise beyond dyadic paradigms to group tasks and interactions. Indeed, some of the effects and psychosocial implications of synchrony are quite possibly magnified rather than diluted through group coordination. Zimmermann & Davidson (2016), for example, showed positive effects on affiliation and on memory and group action video game tasks following manipulation of the synchrony of their verbal behaviour. Unsurprisingly behavioural synchrony between groups of people also leads to greater generosity and subsequent prosocial acts towards co-performers (Reddish, 2013). This effect notably extends to passive 'in-group' members and those primed to be 'non-group' members (Reddish et al., 2016). Thus, it seems that under certain intergroup contexts synchrony can lead to a generalized prosociality (see also Reddish et al., 2014; Tuncgenc & Cohen, 2016). An initially counterintuitive finding was that, in a study of the effects of arbitrarily allocated group membership, entrainment between individuals was most pronounced and secure when participants interacted with a member of a different minimal group (Miles et al., 2010b). Although this would perhaps appear to go against the notion of a general connection between entrainment and social connectedness it is, according to the authors, consistent with a view that "in the absence of any disincentive to affiliate with dissimilar others, interpersonal coordination may act as a medium to support the reduction in intergroup differences and lessen social distance" (Ibid. p. 501)[145]. Finally in a series of interesting experiments Baimel and colleagues (2018) argue that behavioural synchrony promotes 'mentalizing'. Following a task of moving cups in time to recordings of the American national anthem in small groups[146] they found that synchronous conditions increased participant's self-reported tendencies and abilities for considering others' mental states. However, this effect did not afford increased *accuracy* in recognising the mental states of others. Significantly they also presented evidence of increased mental state attribution to those participants with whom they had actively synchronised and decreased mental state attribution to socially non-relevant targets. As summarised by the authors these results "offer some evidence that synchrony tunes our minds to the minds of others, and suppresses the projection of mental states to the world writ large" (Ibid. p. 288). This would certainly appear to be a key mechanism underlying the efficacy of music in supporting group ritual and ceremony across cultures[147]. However, synchrony can focus attention on the shared contents and intentions of *all* social interaction and on contextually relevant cues and meanings[148]. In this regard and in the central processes that support the connections between synchrony, entrainment, and social connection music is, yet again, a particularly structured form rather than being intrinsically unique.

Music as Group Coordination

In this final section of the current chapter I will argue that music is fundamentally a form of cultural ceremony and ritual act. Yet again it is not, I posit, intrinsically unique in these regards. However, it is distinctive in providing an attentional structure that easily facilitates group participation and affords music's efficacies of emotional regulation, intersubjectivity, interactive alignment, and social connectedness to be extended to the group level.

From individual experience to communitas

> *(T)he boundary between individuals and their surroundings, especially others participating in ritual with them, may seem to dissolve… a sense of union is encouraged by the coordination of utterance and movement demanded of congregations in many rituals. To sing with others, to move as they move in the performance of a ritual is not merely to symbolize union. It is in and of itself to reunite in the reproduction of a larger order. Unison does not merely symbolize that order but indicates it and its acceptance. The participants do not simply communicate to each other about that order but commune with each other within it. In sum, the state of communitas experienced in ritual is at once social and experiential. Indeed, the distinction between social and experiential is surrendered, or even erased, in a general feeling of oneness with oneself, with the congregation or with the cosmos. (Rappaport, 1999 p. 220)*

---

[144] However, highly controlled laboratory experiments of simple form of action entrainment have failed to support consistent influences on positive affect (e.g., Reddish et al., 2014; Schachner & Garvin, 2010; Wiltermuth & Heath, 2009). Nevertheless, Mogan and colleagues (2017) maintain that it remains "strong theoretical contender" in naturalistic settings in that even a subtle effect of synchrony on positive affect in constrained laboratory conditions in which affective "confounds" are deliberately minimized, "would be consistent with widely postulated affective mechanisms at work in human rituals" (Ibid p. 14). A recent investigation of large naturalistic groups supports this showing that synchrony and physiological arousal positively affect feelings of cohesion and cooperation (Jackson et al., 2018).

[145] Another possibility is that it reflects a relationship between higher degrees of afforded flexibility of pulse in more securely connected partnerships (see chapter 2).

[146] As there participants were Canadian this was chosen to equate to American participants listening to 'O Canada in the study by Wiltermuth and Heath (2009).

[147] It seems very likely to me that similar findings and conclusions to those found in this section could also result from studying the psychosocial implications of simple tone matching tasks. Extensions of the 'sensorimotor harmonisation' experimental paradigm proposed in chapter 3 could hopefully shed some more light on this in future research.

[148] Even in experimental settings synchrony notably promotes cooperation more powerfully when it is framed as a collective goal. A recent study by Reddish and colleagues (2013) showed that "synchrony combined with shared intentionality leads to greater cooperation than synchrony without shared intentionality or shared intentionality combined with asynchronous movement or vocalising" (p. 1). Synchrony, they posit "interacts with explicitly shared goals to support cooperative interactions" (Ibid).

Where music is mostly clearly distinct from other forms of communicative repertoires and contexts is in cases of group "synchrony" and "harmonization" – "symphonic" or "choral" performances in which all participants structure and align their contribution with regard to a shared musical pulse and/or musical tone/pitch centres. It is crucially important, in my view, to be clear on this description. As discussed in the previous section recent years have witnessed a focus of research on synchronized action with music often being viewed as a pronounced architype of group action synchrony. However, in contrast to group actions such as marching, rowing, or rocking chairs, there is no restriction to produce functionally shared action patterns or even to articulate in shared moments of absolute synchrony. Instead coordination in music is, in the most fundamental sense, structured in reference to mutually and psychologically manifest temporal and frequency-related frameworks. These are undoubtedly and almost ubiquitously reflected in a shared structure of emphasis and overarching dynamic form. However, ultimately these architectural roots can be operational and mutually manifest without being starkly physicalized (e.g. the beat may be entirely implicit) and in the absence of any directly matched actions.

Another critical point to make at this point is that it is, in my opinion, the same specific features of musical pulse and tone that afford a particular framework for the management of personal affective experience and intersubjective interaction that also potentiate these impressive symphonic and/or choral displays and collaborative exchanges across cultures. Nothing additional, in terms of the most basic structuring mechanisms, is required in moving from individual, to dyadic, to increasing group size in music-making. One could even argue that the more challenging it is to achieve secure coordination (e.g. with increasing group size, or rhythmic complexity) the more exclusively we need to rely on these mechanisms and, therefore, the more strictly/inflexibly we adhere to a basic pulse. By focusing participation in relation to mutually manifest pulse and tone we can posit that group size in a musical context can, uniquely, expand without further taxing each individual's attentional load in interaction. This is, of course, a marked difference to other forms of human communication and, in certain key cultural contexts in particular, provides a crucial motivational force for engaging musically. Music's efficacies of emotional regulation, intersubjectivity, interactive alignment, and social connectedness can all, with relative ease, be extended to a group level. In group music-making we are therefore afforded a *structured* and *regulated* form of communitas – one that can be more deliberately and delicately motivated towards a sense of shared experience and time and a convergence of appropriate affective state and attentional focus.

Cultural ceremony

Across cultures, the psychosocial group dynamics of music are consistently at their most potent in conjunction with ritual activity and in the context of large-scale religious ceremony, ritual, and during the creation/reimagining of social and cultural truths (see e.g. Rappaport, 1999; Bohlman, 2002; see also chapter 1). Musical "synchrony" and sacred ceremony mutually enhance their effects on perceptions of shared experience, feelings of "oneness", and facilitate and intensify prosocial behaviors (Fischer et al., 2013)[149]. Similarly to much of the preceding discussion on entrainment, ritual actions increase affiliation (to a greater extent than group activity alone) with in-group members from early childhood onwards (Wen et al., 2016). The psychological literature on rituals supports comparably three primary regulatory functions: regulation of emotions, performance goal states, and social connection (see Hobson et al., 2017 for review). Group rituals and group music-making also can both operate across cultures at relatively low levels of intensity through to being, at times, generative of very high levels of affective arousal – at times the very extremes of transcendent experience[150]. They extend from more general emotional effects of group activity, from a basic and healthy "hive" dynamic (Haidt et al., 2008) or "collective effervescence"[151] (Durkheim, 1912; see also Ehrenreich, 2007). Both "compel participants to feel (or "go through the motions of feeling") emotions appropriate to the purposes of the occasion and ceremonial progression" (Dissanayake, 2006, p. 38). They provide, not a shared experience, but certainly a *convergence* of affective states and attentional focus that are essential in facilitating the operations of cultural and religious ceremony.

Both group ritual action and music are, of course, most commonly associated with religious praxis and belief[152]. They are fundamentally and ideally suited to compliment and support the profoundly personal nature of religious practice with the shared communitas of ritualistic ceremony. Like religion they operate simultaneously at the level of individual experience and collective cultural action and belief. Furthermore, they can be socially facilitative in ritualistic ceremony, by virtue of a relative degree of transposability intentionality, at precisely those moments when verbal accounts could risk conflict or

---

[149] Notably distinguishing between 'exact' synchrony (e.g. Buddhist *unison* chanting) and 'complementary' synchrony (e.g. Brazilian drumming or choral singing [*in parts*]) this project studied nine natural rituals. Their statistical analysis suggested a "proximate mechanism by which (a) synchronous actions (b) enhance feelings of oneness (entitativity) to (c) intensify sacred values, thus (d) increasing prosocial behaviors" (Ibid. p. 123).

[150] In both cases, examples are of course as numerous and varied as cultures (and subcultures) across time and geographical divides. The social-affective dimensions of a few selected naturally occurring religious rituals have notably and vividly been portrayed in recent systematic studies ethnographic records (e.g. Bulbulia et al., 2013; Fischer et al., 2014; Xygalatas et al., 2013). While well-researched descriptions of, for example, intense performances with positive valence among the BaYaka Pygmies of the Central African Republic and of the highly intense rituals, singing and trancing found in charismatic Pentecostal Christian congregations are presented in Lewis (2013) and Anderson (2013) respectively. Trehub and colleagues (2015) point out that it may be significant to an evolutionary perspective that the pleasure and social consequences are amplified, when music is experienced with family or peers (e.g. Boer and Abubakar, 2014).

[151] Durkheim famously described this phrase as follows: "Once the individuals are gathered together, a sort of electricity is generated from their closeness and quickly launches them to an extraordinary height of exaltation" (Ibid. p. 217)

[152] Indeed, in some cultures (e.g. Candomble´ Afro-Brazilian religion), music is so ubiquitously connected with religion that the same term covers what we would recognize as music and religion (see Bohlman, 2002).

paradox (Cross 2009; see Chapter 1). Discussing music specifically Bohlman (2002, p. 5) succinctly describes that "[e]xamining music cross-culturally, we recognise that religious meaning accrues to music in many ways. Music may serve as a vehicle for shaping the voice of a deity; it may demarcate time so that it is more meaningful for the performance of ritual; music may provide one of many ornaments that make religious practice more attractive" (emphasis added). It seems highly plausible that the extension of time and attentional emphasis afforded through the specific features of music described in this thesis are also operational in supporting a sustained liminal stage in ritual process and the emergence of a new social reality (Van Gennep, 1960; Turner, 2017).

Music as ritual

"[R]itual is unique in at once establishing conventions, that is to say enunciating and accepting them, and in insulating them from usage. In both enunciating conventions and accepting them, it contains within itself not simply a symbolic representation of social contract, but a consummation of social contract. As such, ritual ... is the basic social act." (Rappaport 1979b, p. 197, emphasis in the original)

Crucially music should not be understood as merely playing an attentional and/or affective supportive role in ceremony. Instead music is itself a form of ritual symbol and act. Rappaport distinguished between ritual and drama in terms of the former being intended towards a causal impact on the world. His understanding was that, whereas an actor's role only lasts as long as the performance, by participating in ritual a participant is involved functionally in "the enduring order that their participation brings into being" (p. 136). I would however support the view suggested in a pertinent paper by Morley (2009) that "there is no such thing as pure (polar) ritual or pure (polar) performance (by Rappaport's criteria) and that the continuum is one only of varying relative proportions of the two, the ends of which are not absolutes" (p. 163). He continues that "it is impossible to absolutely separate the two on the basis of the intended temporality of effects of the activity on the world, or on the temporality of the personae adopted. It is also impossible to separate the two on the basis of the extent to which they effect change through appeal to a supernatural agent rather than through other humans" (Ibid). Music can, of course operate at both extreme ends of the spectrum. For example at the Kaiko festival of the Maring of Papua New Guinea (Rappaport 1999) the dance signifies a commitment to the hosting group. To dance is more than a symbol; it is an indexical indication – the pledge enacted. Morley, following Rappaport, concludes that "[a] fundamental characteristic of ritual here is its ability to bring about a conventional state of affairs – to affirm, transform, or bring into being a particular natural order – that is, it is meant to affect the world" (Ibid p. 162).

Music can also be understood as a form of ritual in an even more fundamental sense. "Ritualized" actions - highly stereotyped symbols (Huxley, 1914) - are common in fish and birds (Tinbergen, 1952; Lorenz 1950; 1966), and in mammals including nonhuman primates (Hinde 1974; Moynihan 1956). These are typically components of a behaviour "that occurs as part of normal, everyday, instrumental activity—such as preening, nestbuilding, preparing to fly, or caring for young—are, as it were, "selected" or taken out of context, "ritualized," and used to signal an entirely different motivation—usually an attitude or intention that may then influence (affect or manipulate) the behavior of another animal" (Dissanayake, 2006 p. 4). Following Smith (1977), Watanabe and Smuts (1999) propose three essential aspects of animal ritualization. "First, a behavior that originally served a specific utilitarian function appears in a new social context. Second, in this new context, the behavior typically becomes simplified, exaggerated, and stereotyped. Third, the new function of the behavior has less to do with achieving immediate, instrumental ends than with signaling an attitude or intention that, in turn, may influence the behavior of another individual" (p. 100). These ceremonial displays are also "ritualized" in the sense that they attract attention and influence emotions and behaviour, and also prescribe the scope of possible behavior and provide a mutually recognized, agreed, and formalized framework for interaction. They can, furthermore, "literally embody in communal participation the mutual coordination they presuppose" (Dissanayake, 2006 p. 5). In a series of papers Dissanayake (e.g. 1979; 2001; 2015) has argued that human ritual ceremony and its associated culturally-constructed arts (including music) extend from such instinctive biological displays and that in certain contexts, therefore, the "inclination to make ordinary behavior extraordinary" (2006, p. 6) is part of our biological nature. She notes, in support, "that a large proportion of the distinctive recombined components of human ritual ceremonies resemble (or in fact are) what we call the arts—dance and mime, poetic language, visual display, and "music" (song)" and that "one can view ceremonial and other arts as ordinary behavior (i.e., ordinary bodily and facial movements, ordinary speech, utilization of ordinary objects and surroundings, and ordinary prosodic vocalizations) made extraordinary through essentially the same operations or procedures as in animal ritualizations: formalization (stereotypy), repetition, exaggeration, and elaborations of various kinds" (Ibid, p. 6). Understanding music as a form of ritual in this way makes perfect sense in terms of the central arguments in this thesis. Music is motivated by our most central human desires and drives to communicate, experience, express and regulate emotions, to connect, to share goals and/or align with others,. However there is an additional overarching motivation to structure these in a particular form – a form of formalization, repetition, exaggeration, and elaboration. In the case of music, this form and its corresponding motivation structures human social and affective drives within an architectural form rooted in configurations of musical pulse and tone.

Chapter Summary

Music is intrinsically motivated by our most central human motivations of emotional experience, expression, and regulation; intersubjectivity; social alignment; cultural belonging, and communitas. It is a prominent component in our suite of behaviours directed towards achieving hedonic and eudaimonic well-being through meaningful personal reflection and socio-intentional connection; it is culturally generative and reflective, an artistic expression of individual creativity, character and of cultural belonging. These motivations and their dynamic correlates are, however, not specifically enacted in music; they all emerge early in ontogeny and are manifest in many active and social forms in the course of human development. They are instead given and afforded particular space in music – a space that constantly binds from one perceptual present to the next; offers a constant renewal of attentional focus; provides a degree of predictable continuity; and can be interactively manifest in groups. Therefore the distinct motivation for music lies in giving our central affective and socio-intentional drives extended phenomenological space, stability, and a degree of abstraction, intensity, and meaning and in providing a framework for interpersonal and group interaction and for ritual action.

Music is most fundamentally motivated towards providing an extended phenomological experience, a particular attentional structure for the management of individual experience, and a coordinative socio-intentional framework for interaction affording group synchrony and an intersubjective, convergence of affect.

Section III – Summary and Conclusions

The Narrow Faculty for Music

Musical pulse, musical tone and musical motivation are respectively embedded in and have developed from our most foundational and central kinaesthetic, vocal, affective, and socio-intentional drives and capacities. In terms of fundamental temporal and frequency based processing there seems to have been a profound, if not categorically unique shift, towards relative processing in our species. Analysis of the specificities of a musical pulse and tone both focused in on correction mechanisms that tie the immediate dynamic past with the present moment and expected moments of future attentional emphasis. Both appeared to provide an awareness of the architectural framework and, crucially, to be subject to a degree of volitional control. Accordingly the specific nature of musical motivation was primarily described as a drive towards extending and intensifying experience and in providing an attentionally rooted structure for managing individual practice as well as joint action and alignment. In comparative analysis within broad frameworks of human and animal communication, meaning, and emotion, the notion was thus put forward that the constituent features of the narrow faculty for music are proximally functional in providing an extended coordinative framework for meaningful experience, delimited social interaction and the efficient and appropriate regulation of affect individually and within groups. Effectively, by directing and sustaining attention to the afferent and efferent significances of interaction - their dynamic affective-motivational processes - and making them available to our awareness, the constituent parts of our narrow faculty for music can enhance and stabilize their effects.

More specific and extensive summaries of the individual chapters are available in the final paragraphs of each of the chapters respectively. It is worth reiterating, however, why the thesis has focused in on the argument that features that are specific to music provide a specifically *attentional* and motivational framework for the management of individual experience and social interaction. This follows directly from the extensive available evidence that the most fundamental architectural foundation of temporal structure in music (the 'beat') is, in essence, an internally generated attentional pulse. It is part of wider timing mechanisms and our capacities to dynamically attend to events in time and to interact with the continuously moving physical and social world. Future-directed expectations allow us to parsimoniously allocate attentional resouces in all areas of life and are a crucial social affordance in enabling all forms of interpersonal interaction. However, my analysis suggests that a musical pulse is distinct, not in an absolute sense, but in the particular nature of sustained attention; its continually operational period correction mechanisms; a corresponding degree of volitional control; and in the awareness of the underlying temporal beat.

Attentional structure was also shown to be crucial in describing the specificities of musical tone. Although the link is perhaps less direct than with pulse, it is again correction mechanisms that afford sustained utterance and attention, a high degree of volitional control, and an awareness of the architectural framework that appear to be distinct to the context of music. Crucially, correction mechanisms of pulse and tone are inherently a form of sustained attentional structure. In both cases they involve holding in mind/attention the events of the immediate past, the current moment ("now"), and future expectation and continously linking them into an extended perceptual present. Finally, attention was again central in describing the specificities of musical motivation in that the specific motivation for music was described as an impetus to give our most central socio-intentional and affective drives extended phenomenological space within the attentional structures of configurations of a musical pulse and/or musical tone as described in chapters 2 and 3.

Evolutionary Implications

It was not the main purpose of this thesis to enter into debates on the relative weight that should be attributed to the various putative evolutionary trajectories and rationales in the literature. A full generic capacity for music incorporates skills and mechanisms that were putatively around in the earliest jawed vertebrates some 500 million years ago (Chase, 2001). Hence, one of my principal arguments is that to debate the adaptive functionality of music or its evolutionary status as an adaptation without first defining the matter for debate and identifying putatively its unique features is necessarily incomplete. We risk falsely attributing functionalities to "music" that are in fact more precisely attributable to broader categories of relevance. Nevertheless the analysis undertaken does of course support some hypotheses more than others and suggests views that could be extended in future research and publications.

First of all, having focused on the specificities of music in relative processing, correction mechanisms, awareness, and intention, I argue that an exaptive origin for music is entirely plausible. The capacities that allow us to structure our individual actions, attention, and interactions in musical form could quite conceivably have arisen as exaptations upon a socially-driven move towards relational empathy (Trevarthen & Malloch, 2018); mimetic interaction (Donald, 1991); enhancements in the abilities to sustain attention to events—active or imaginary; to volitionally control actions (Norman & Shallice, 1980); to rehearse events in short-term working memory (Baddeley, 2003; see also Olson & Kawamar 1999); and/or the ability and/or motivation to manage the cognitive environment of others in communication (Sperber & Wilson, 1986). There is no definite reason to assume a biological adaptive origin for music[153]. Secondly, I suggest that the

---

[153] As noted in Bispham (2006) the evolutionary function of the specificities of music may also have been dependent on other species-specific characteristics. For example, increased infant altriciality, expanding group size, and increased sociality in the hominin lineage (Dunbar, 1993; Joffe, 1997) may have uniquely provided selection pressures toward the use of affective and contingent modes of communication, providing a necessary stepping-stone to music. (p. 131).

description of the specific features of music in this thesis argues against the need for a group selection model for music. The features that allow a particular framework for attention and the management of individual experience are precisely those that afford group synchrony and harmony. It is therefore be possible to posit an evolutionary trajectory for group music-making without necessarily encountering some of the problematic concerns of group-level selection models. Contrary to Dissanayake's (2006) view, music may therefore not be "the first cognitive adaptation that is not completely explainable by individual selection mechanisms (Brown, 2000)" and is not necessarily "an important arrow in the quiver of advocates of group selection arguments" (p. 6).

Modern-day Implications

The description of the narrow faculty for music also offers some potential insight as to generalizable mechanisms operating in the often rather nebulously understood "power of music" in individual therapy and cultural music-making. Crucially, the central mechanisms of expressing and/or regulating emotions, of relating to others, and our drives towards feelings of belonging and attachment are not viewed here as being in any way distinct in music. Instead they are given particular structure and substance in music. I argue therefore that progress will come from appreciating that the specific efficacies of music need to be understood in terms of the transformative properties and overarching affordances of its structural roots and the corresponding architecture of attention involved. In emotional music therapy, for example, we need not posit any disconnect from more widespread regulatory effects of interpersonal interaction. Rather, the patient is given protracted time to realise some of the testing intricacies of communicative intent in abstracted form. He/she may benefit from an attentional strategy that guides, stabilises and allows a degree of control over the constant flux of emotional furour and challenge. More generally, and without seeking to diminish the affective nature of musical experience, I suggest that the course of argument in this thesis advocates that the *specific* features of our capacity for music need to be understood primarily in terms of issues of attention, focus and concentration. These features extend enacted experience, enhance affect, and intensify meaning of the particular actions and interactions of a given musical style and/or individual practice.

Final Comments

I submit that, whether or not the reader agrees with all parts of my preceding argument, discussions of music in evolution as well as our support for the continuing value of music in modern-day society need to be focused on and structured in accordance to an understanding of the specificities of our human capacity for music – the narrow faculty for music. Finally I would like to reiterate that the central propositions of this thesis - positing the specificities of music in its seemingly most basic firmament of a pulse, tone, and motivation - should not be understood as presenting a reductionist account of music or as a position that discussion on music should ever be exclusively framed in these terms. Rather they constitute a potentiating space, a particular experience of action in time that affords and supports the extraordinary flexibility of dynamic form and function across cultures. They afford an extended moment in which a virtually infinite variety of individual creativity, emotional reality and regulation, social connection, group coordination, and cultural history and reflection can be given stability and an intensity of focus, memory, emotion, and meaning.

The constituent parts of the narrow faculty for music – musical pulse, musical tone, and musical motivation – provide a sustained attentional framework for managing personal experience and a coordinative structure for interpersonal interaction. They extend a phenomenological present thus affording a continuing focus, and intensity of affect, stability, regulation, memory, and meaning.

Acknowledgments

First of all, I would like to express my sincere gratitude to my PhD supervisor Professor Ian Cross for his continuous support, kindness, generosity, and academic expertise. I would also like to thank my two examiners Professor Elizabeth Tolbert and Dr Robert Attenborough for their time, efforts, and for insightful comments and suggestions. Further thanks go to many friends and colleagues at the Centre for Music and Science, all of whom offered valuable contributions in seminars and discussions and also helped me to remain reasonably sane: Neta Spiro (thank you also for proof-reading), Tommi Himberg, Joel Swaine, Elizabeth Blake, Satinder Gill, Nick Collins, Matthew Woolhouse, Nicki Moran, Justin London, and Jessica Grahn. Particular thanks are also owed to Prof Bill Thompson for his empathy, inspiration, and compassion. Finally, thanks to my mother and sister who proofread large sections of the thesis and much more importantly to them both and my father for years of encouragement and familial support.

References

Abdelhak, Kalatzis, et al. (1997). "Clustering of mutations responsible for branchio-oto-renal (BOR) syndrome in the eyes absent homologous region (eyaHR) or EYA1." Human Molecular Genetics **6**: 2247-2255.

Ackermann, H., S. R. Hage, et al. (2014). "Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective." Behavioral and Brain Sciences **37**(6): 529.

Adachi, M. and S. E. Trehub (1998). "Children's expression of emotion in song." Psychology of Music **26**(2): 133-153.

Adler, C. H. and J. E. Ahlskog (2000). Parkinson’s disease and movement disorders: diagnosis and treatment guidelines for the practicing physician, Springer Science & Business Media.

Aiello, L. and R. Dunbar (1993). "Neocortex size, group size, and the evolution of language." Current Anthropology **34**: 184-193.

Aiello, L. C., T. Wynn, et al. (2010). "Working memory: Beyond language and symbolism." Current Anthropology **51**.

Aitken, P. G. and W. A. Wilson (1979). "Discriminative vocal conditioning in rhesus monkeys: Evidence for volitional control?" Brain and language **8**(2): 227-240.

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

Albouy, P., J. Mattout, et al. (2013). "Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex." Brain **136**(5): 1639-1661.

Albouy, P., J. Mattout, et al. (2015). "Altered retrieval of melodic information in congenital amusia: insights from dynamic causal modeling of MEG data." Frontiers in human neuroscience **9**: 20.

Alcock, K. J., R. E. Passingham, et al. (2000). "Pitch and timing abilities in inherited speech and language impairment." Brain and language **75**(1): 34-46.

Alcorta, C. and R. Sosis (2005). "Ritual, emotion, and sacred symbols." Human Nature **16**(4): 323-359.

Alemseged, Z., F. Spoor, et al. (2006). "A juvenile early hominin skeleton from Dikika, Ethiopia." Nature **443**(21): 296-301.

Allen, D. (1967). "Octave discriminability of musical and non-musical subjects." Psychonomic Science **7**(12): 421-422.

Allman, M. J., S. Teki, et al. (2014). "Properties of the internal clock: first-and second-order principles of subjective time." Annual review of psychology **65**: 743-771.

Allsop, J. S., T. Vaitkus, et al. (2016). "Coordination and collective performance: cooperative goals boost interpersonal synchrony and task outcomes." Frontiers in psychology **7**: 1462.

Altenmueller, E., K. Schuermann, et al. (2002). "Hits to the left, flops to the right: different emotions during listening to music are reflected in cortical lateralization patterns." Neuropsychologia **40**: 2242-2256.

Altenmüller, E., J. Marco-Pallares, et al. (2009). "Neural Reorganization Underlies Improvement in Stroke-induced Motor Dysfunction by Music-supported Therapy." Annals of the New York Academy of sciences **1169**(1): 395-405.

Altenmüller, E. and G. Schlaug (2013). "Neurobiological aspects of neurologic music therapy." Music and medicine **5**(4): 210-216.

Amador, A. and D. Margoliash (2013). "A mechanism for frequency modulation in songbirds shared with humans." Journal of Neuroscience **33**(27): 11136-11144.

Amir, O., N. Amir, et al. (2003). "The effect of superior auditory skills on vocal accuracy." The Journal of the Acoustical Society of America **113**(2): 1102-1108.

Anderson, A. H. (2013). To the ends of the Earth: Pentecostalism and the transformation of world Christianity, Oxford University Press.

Ansdell, G. and M. Pavlicevic (2005). "Musical companionship, musical community. Music therapy and the process and value of musical communication." In D. Miell, R. Macdonald, D. Harreaves (Eds.) Musical communication: 193-213.

Araya-Salas, M. (2012). "Is birdsong music? Evaluating harmonic intervals in songs of a Neotropical songbird." Animal Behaviour **84**(2): 309-313.

Arcadi, A. C., D. Robert, et al. (1998). "Buttress drumming by wild chimpanzees: Temporal patterning, phrase integration into loud calls, and preliminary evidence for individual distinctiveness." Primates **39**(4): 505-518.

Arcadi, A. C., D. Robert, et al. (2004). "A comparison of buttress drumming by male chimpanzees from two populations." Primates **45**(2): 135-139.

Arensberg, P., L. Schepartz, et al. (1990). "A reappraisal of the anatomical basis for speech in middle Paleolithic hominids." American Journal of Physical Anthropology **83**: 137-146.

Arom, S. (1991). African polyphony and polyrhythm: Musical structure and methodology. Cambridge, Cambridge University Press.

Arom, S., N. Fernando, et al. (2007). An innovative method for the study of african musical scales: cognitive and technical aspects. Proceedings of the 4th sound and music computing conference.

Arom, S. and J. Khalfa (1998). "Une raison en acte: Pensee formelle et systematique musicale dans les societes de tradition orale." Revue de Musicologie **84**: 5-17.

Arsuaga, J., I. Martinez, et al. (1997). "Sima de los Huesos (Sierra de Atapuerca, Spain): the site." Journal of Human Evolution **33**: 109-127.

Arvaniti, A. (2009). "Rhythm, timing and the timing of rhythm." Phonetica **66**(1-2): 46-63.

Aschersleben, G. (2002). "Temporal control of movements in sensorimotor synchronization." Brain and cognition **48**(1): 66-79.

Aschersleben, G., J. Gehrke, et al. (2001). "Tapping with peripheral nerve block." Experimental brain research **136**(3): 331-339.

Aschersleben, G. and W. Prinz (1995). "Synchronizing actions with events: The role of sensory information." Attention, Perception, & Psychophysics **57**(3): 305-317.

Aschersleben, G. and W. Prinz (1997). "Delayed auditory feedback in synchronization." Journal of Motor behavior **29**(1): 35-46.

Aschersleben, G., P. Stenneken, et al. (2002). Timing mechanisms in sensorimotor synchronization. Common mechanisms in perception and action, Oxford University Press**:** 227-244.

Ashby, W. R. (1961). An introduction to cybernetics, Chapman & Hall Ltd.

Ashton–James, C., R. B. Van Baaren, et al. (2007). "Mimicry and me: The impact of mimicry on self–construal." Social cognition **25**(4): 518-535.

Assisi, C. G., V. K. Jirsa, et al. (2005). "Dynamics of multifrequency coordination using parametric driving: theory and experiment." Biological cybernetics **93**(1): 6-21.

Astafiev, S. V., G. L. Shulman, et al. (2003). "Functional organization of human intraparietal and frontal cortex for attending, looking, and pointing." Journal of Neuroscience **23**(11): 4689-4699.

Atran, S. (2002). In gods we trust: The evolutionary landscape of religion. Oxford, Oxford University Press.

Auer, P., E. Couper-Kuhlen, et al. (1999). Language in time: The rhythm and tempo of spoken language. New York, Oxford University Press.

Avanzini, G., L. Lopez, et al. (2006). "The neurosciences and music II: From perception to performance (special issue)." Annals of the New York Academy of sciences **1060**.

Avanzini, P., R. O. Abdollahi, et al. (2016). "Four-dimensional maps of the human somatosensory system." Proceedings of the National Academy of Sciences **113**(13): E1936-E1943.

Avril, C. (1963). "Kehlkopf und Kehlsack des Schimpansen (Pan troglodytes)." Gegenbaurs Morphologie Jahrbuch **195**: 74-129.

Ayala, F. and A. Escalante (1996). "The evolution of human populations: a molecular perspective." Molecular Phylogeny and Evolution **5**(188-201).

Ayotte, J., I. Peretz, et al. (2002). "Congenital amusia: a group study of adults afflicted with a music-specific disorder." Brain **125**(2): 238-251.

Bachorowski, J. and M. Owren (1995). "Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context." Psychological Science **6**: 219-224.

Bachorowski, J. and M. Owren (2003). "Sounds of emotion: Production and percettion of affect-related vocal acoustics." Annals of the New York Academy of Sciences **1000**: 244-265.

Baddeley, A. (2003). "Working memory: looking back and looking forward." Nature reviews. Neuroscience **4**(10): 829.

Bahrick, L. E. and R. Lickliter (2004). "Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis." Cognitive, Affective, & Behavioral Neuroscience **4**(2): 137-147.

Bailey, B. and J. Davidson (2005). "Effects of group singing and performance for marginalized and middle-class singers." Psychology of Music **33**(3): 269-303.

Bailey, B. A. and J. W. Davidson (2005). "Effects of group singing and performance for marginalized and middle-class singers." Psychology of Music **33**(3): 269-303.

Baillargeon, R., R. M. Scott, et al. (2010). "False-belief understanding in infants." Trends in cognitive sciences **14**(3): 110-118.

Baily, J. (1985). Music structure and human movement. Musical structure and cognition. P. Howell, I. Cross and R. West. London, Academic Press**:** 237-258.

Baily, J. (2005). So near, so far: Kabul's music in exile. Ethnomusicology Forum, Taylor & Francis.

Baimel, A., S. A. Birch, et al. (2018). "Coordinating bodies and minds: Behavioral synchrony fosters mentalizing." Journal of Experimental Social Psychology **74**: 281-290.

Baimel, A., R. L. Severson, et al. (2015). "Enhancing "theory of mind" through behavioral synchrony." Frontiers in psychology **6**: 870.

Bakeman, R. and L. B. Adamson (1984). "Coordinating attention to people and objects in mother-infant and peer-infant interaction." Child development: 1278-1289.

Baker, F.. Gleadhill, et al. (2007). "Music therapy and emotional exploration: Exposing substance abuse clients to the experiences of non-drug-induced emotions." The Arts in Psychotherapy **34**(4): 321-330.

Baldwin, D. A., J. A. Baird, et al. (2001). "Infants parse dynamic action." Child development **72**(3): 708-717.

Balkwill, L. and W. F. Thompson (1999). "A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues." Music Perception **17**: 43-64.

Balkwill, L. & W. F. Thompson, et al. (2004). "Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners." Japanese Psychological Research **46**(4): 337-349.

Balter, M. (2004). "Seeking the key to music." Science **306**: 1120-1122.

Baluch, F. and L. Itti (2011). "Mechanisms of top-down attention." Trends in neurosciences **34**(4): 210-224.

Balzano, G. J. (1980). "The group-theoretic description of 12-fold and microtonal pitch systems." Computer music journal **4**(4): 66-84.

Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In Music, mind, and brain, Springer**:** 321-351.

Banfield, J. F., C. L. Wyland, et al. (2004). "The cognitive neuroscience of self-regulation."In K. Vohs & V. Baumeister (Eds) Handbook of self-regulation: Research, theory, and applications **62-83**.

Bannan, N. (2012). "Harmony and its role in human evolution." In Music, language, and human evolution: 288-339.

Bard, K. A. and J. Vauclair (1984). "The communicative context of object manipulation in ape and human adult-infant pairs." Journal of Human Evolution **13**(2): 181-190.

Barkow, J., L. Cosmides, et al. (1992). The adapted mind. Oxford, Oxford University Press.

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

Baron-Cohen, S. (1997). Mindblindness: An essay on autism and theory of mind, MIT press.

Barrett, J. and A. S. Fleming (2011). "Annual research review: All mothers are not created equal: Neural and psychobiological perspectives on mothering and the importance of individual differences." Journal of Child psychology and Psychiatry **52**(4): 368-397.

Barrow, J. (1995). The artful universe. Oxford, Clarendon Press.

Bartolo, R. and H. Merchant (2015). "β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey." Journal of Neuroscience **35**(11): 4635-4640.

Bartolo, R., L. Prado, et al. (2014). "Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping." Journal of Neuroscience **34**(11): 3910-3923.

Basso, E. B. (1985). A musical view of the universe: Kalapalo myth and ritual performances, University of Pennsylvania Press.

Bates, E., S. D'Amico, et al. (2003). "Timed picture naming in seven languages." Psychonomic bulletin & review **10**(2): 344-380.

Baumeister, R. F. and M. R. Leary (1995). "The need to belong: desire for interpersonal attachments as a fundamental human motivation." Psychological bulletin **117**(3): 497.

Becker, J. (2001). Anthropological perspectives on music and emotion. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press.

Beckman, M. (1982). "Segment duration and the 'mora'in Japanese." Phonetica **39**(2-3): 113-135.

Beebe, B., J. Jaffe, et al. (1992). A dyadic systems view of communication. Relational perspectives in psychoanalysis. N. Skolnick and S. Warshaw. Hillsdale, NJ, Analytic Press**:** 61-81.

Beebe, B., S. Knoblauch, et al. (2005). Forms of intersubjectivity in infant research and adult treatment, Other Press.

Beedie, C., P. Terry, et al. (2005). "Distinctions between emotion and mood." Cognition & Emotion **19**(6): 847-878.

Behne, K. (1997). The development of 'musikerleben' in adolescence: How and why young people listen to music. Perception and cognition of msuic. I. Deliege and J. Sloboda. Hove, UK, Psychology Press**:** 143-159.

Behne, T., M. Carpenter, et al. (2005). "Unwilling versus unable: infants' understanding of intentional action." Developmental Psychology **41**(2): 328.

Behroozmand, R., H. Liu, et al. (2011). "Time-dependent neural processing of auditory feedback during voice pitch error detection." Journal of Cognitive Neuroscience **23**(5): 1205-1217.

Behroozmand, R., S. Sangtian, et al. (2016). "A temporal predictive code for voice motor control: Evidence from ERP and behavioral responses to pitch-shifted auditory feedback." Brain research **1636**: 1-12.

Belfi, A. M., E. Evans, et al. (2017). "Musical anhedonia after focal brain damage." Neuropsychologia **97**: 29-37.

Belin, P., S. Fecteau, et al. (2004). "Thinking the voice: Neural correlates of voice perception." Cognitive Science **8**(3): 129-136.

Belyk, M. and S. Brown (2017). "The origins of the vocal brain in humans." Neuroscience & Biobehavioral Reviews.

Bendor, D. and X. Wang (2005). "The neuronal representation of pitch in primate auditory cortex." Nature **436**(7054): 1161.

Bendor, D. and X. Wang (2006). "Cortical representations of pitch in monkeys and humans." Current opinion in neurobiology **16**(4): 391-399.

Bendor, D. and X. Wang (2010). "Neural coding of periodicity in marmoset auditory cortex." Journal of neurophysiology **103**(4): 1809.

Bengtsson, S. L., F. Ullén, et al. (2009). "Listening to rhythms activates motor and premotor cortices." Cortex **45**(1): 62-71.

Benson, D. (1985). Language in the left hemisphere. The dual brain. D. Benson and E. Zaidel. New York, Guildford Press**:** 193-203.

Benzon, W. (2001). Beethoven's anvil: Music in mind and culture. Oxford, Oxford University Press.

Bergstrom, T., A. Josefsson, et al. (1998). "Recent origin of HLA-DRB1 alleles and implications for human evolution." Nature Genetics **18**: 237-242.

Bermudez de Castro, J., J. Arsuaga, et al. (1997). "A Hominid from the Lower Pleistocene of Atapuerca, Spain: Possible Ancestor to Neandertals and Modern Humans " Science **276**(5317): 1392-1395.

Bernstein, L. (1976). The unanswered question: Six talks at Harvard. Cambridge, Massachusetts, Harvard University Press.

Bertelson, P. and G. Aschersleben (2003). "Temporal ventriloquism: crossmodal interaction on the time dimension: 1. evidence from auditory–visual temporal order judgment." International Journal of Psychophysiology **50**(1): 147-155.

Bertenthal, B. I. (1996). "Origins and early development of perception, action, and representation." Annual review of psychology **47**(1): 431-459.

Bertinetto, P. M. (1989). "Reflections on the dichotomy 'stress' vs.'syllable-timing'." Revue de phonétique appliquée **91**(93): 99-130.

Besson, M. and D. Schon (2001). "Comparison between language and music." Annals of the New York Academy of Sciences **930**: 232-258.

Bharucha, J. (1994). Tonality and expectation. Musical perceptions. R. Aiello and J. Sloboda. New York, Oxford University Press**:** 213-239.

Białuńska, A., S. Dalla Bella, et al. (2011). "Increasing stimulus intensity does not affect sensorimotor synchronization." Psychological Research **75**(1): 43-53.

Biasutti, M. (2013). "Orchestra rehearsal strategies: conductor and performer views." Musicae Scientiae **17**(1): 57-71.

Bidelman, G. M. (2013). "The role of the auditory brainstem in processing musically relevant pitch." Frontiers in psychology **4**.

Bidelman, G. M. and J. Grall (2014). "Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex." Neuroimage **101**: 204-214.

Bidelman, G. M. and M. G. Heinz (2011). "Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing a." The Journal of the Acoustical Society of America **130**(3): 1488-1502.

Bidelman, G. M. and A. Krishnan (2009). "Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem." Journal of Neuroscience **29**(42): 13165-13171.

Bijsterbosch, J. D., K.-H. Lee, et al. (2011). "Continuous theta burst stimulation over the left pre-motor cortex affects sensorimotor timing accuracy and supraliminal error correction." Brain research **1410**: 101-111.

Bijsterbosch, J. D., K.-H. Lee, et al. (2011). "The role of the cerebellum in sub-and supraliminal error correction during sensorimotor synchronization: evidence from fMRI and TMS." Journal of cognitive neuroscience **23**(5): 1100-1112.

Billon, M., C. Bard, et al. (1996). "Simultaneity of two effectors in synchronization with a periodic external signal." Human Movement Science **15**(1): 25-38.

Bilney, B., M. E. Morris, et al. (2005). "Evidence for a disorder of locomotor timing in Huntington's disease." Movement disorders **20**(1): 51-57.

Bindra, D. (1969). "A unified interpretation of emotion and motivation." Annals of the New York Academy of sciences **159**(1): 1071-1083.

Bispham, J. (2002). Infant-directed Lullaby - A Cross-Cultural Genre. Essay submitted in partial recognition of MPhil in Musicology, University of Cambridge, Faculty of Music

Bispham, J. (2003). An evolutionary perspective on the human skill of interpersonal musical entrainment. Music Department. Cambridge, Cambridge University.

Bispham, J. (2004). "Bridging the Gaps–Music as Biocultural Phenomenon." ESEM Counterpoint **1**: 78-81.

Bispham, J. (2006). "“Music” means nothing if we don't know what it means." Journal of Human Evolution **50**(5): 587-593.

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

Bispham, J. (2007). "Music as Socio-Affective Confluential Communication? Response to Graham." Music Perception: An Interdisciplinary Journal **25**(2): 169-170.

Bispham, J. C. (2009). "Music's “design features”: Musical motivation, musical pulse, and musical pitch." Musicae Scientiae **13**(2_suppl): 41-61.

Bispham, J. C. (2013). "How Musical is Man? An Evolutionary Perspective." Sound Musicianship: Understanding the Crafts of Music.

Blacking, J. (1969). "The value of music in human experience." Yearbook of the International Folk Music Council **1**: 33-71.

Blacking, J. (1973). How musical is man? Seattle, University of Washington Press.

Blacking, J. (1995). Music, culture, and experience: Selected papers of John Blacking, University of Chicago Press.

Blackwell, H. R. and H. Schlosberg (1943). "Octave generalization, pitch discrimination, and loudness thresholds in the white rat." Journal of experimental psychology **33**(5): 407.

Blake, E. (2006). Assessing lithophone technology for evidence of intentional sound production. Department of Archaeology. Cambridge, University of Cambridge, UK. **MPhil**.

Bless, H. and J. Forgas, Eds. (2000). The message within. The role of subjective experience in social cognition and behaviour. Hove, UK, Psychology Press.

Block, R. A. (2014). Cognitive models of psychological time, Psychology Press.

Block, R. A. and R. P. Gruber (2014). "Time perception, attention, and memory: a selective review." Acta psychologica **149**: 129-133.

Blood, A. J. and R. J. Zatorre (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.

Blumstein, D. and K. Armitage (1997). "Alarm-calling in yellow-bellied marmots. I. The meaning of situationally variable alarm calls." Animal Behaviour **53**: 143-171.

Boer, D. and A. Abubakar (2014). "Music listening in families and peer groups: benefits for young people's social cohesion and emotional well-being across four cultures." Frontiers in psychology **5**: 392.

Boesch, C. (2005). "Joint cooperative hunting among wild chimpanzees: Taking natural observations seriously." Behavioral and Brain Sciences **28**(5): 692-693.

Bögels, S., L. Magyari, et al. (2015). "Neural signatures of response planning occur midway through an incoming question in conversation." Scientific reports **5**.

Bögels, S. and F. Torreira (2015). "Listeners use intonational phrase boundaries to project turn ends in spoken interaction." Journal of phonetics **52**: 46-57.

Bohlman, P. (2000). Ethnomusicology and music sociology. Musicology and sister disciplines. D. Greer. Oxford, Oxford University Press**:** 288-298.

Bohlman, P. V. (2002). World music: A very short introduction, OUP Oxford.

Bolhuis, J. J. and C. D. Wynne (2009). "Can evolution explain how minds work?" Nature **458**(7240): 832-833.

Bolinger, D. (1964). Intonation across languages. Universals in human language. J. Greenberg, C. Ferguson and E. Moravcsik. Stanford, CA, Stanford University Press**:** 313-318.

Bolinger, D. (1978). Intonation across languages. Universals in human language: Volume 2. phonology. J. Greenberg, C. Ferguson and E. Moravcsik. Pao Alto, CA, Stanford University Press**:** 472-524.

Bolt, L. (2007). "Sexual selection and the origins of music." Cambridge Archaeological Journal.

Bolton, T. L. (1894). "Rhythm." The american journal of psychology **6**(2): 145-238.

Boone, B. (2003). "Pitch synchrony in Live with Regis and Kathie Lee dialogue." 2003 Hawaii.

Boone, B., Basiletti, S. et al. (2007). “An exploratory acoustic analysis of musical patterns of pitch variation in dyadic conversation,” Poster presented at the 28th Annual Central California Research Symposium, Fresno, CA.

Botha, R. (2009). "On musilanguage/“Hmmmmm” as an evolutionary precursor to language." Language & communication **29**(1): 61-76.

Bowlby, J. (1969). Attachment and Loss: Attachment; John Bowlby, Basic Books.

Bowlby, J. (1973). Attachment and loss: Separation (vol. 2), New York: Basic Books.

Bowles, S. and H. Gintis (1998). The evolution of strong reciprocity. Santa Fe, N.M, Santa Fe Institute.

Bowles, S. and H. Gintis (1998). "The moral economy of community: structured populations and the evolution of pro-social norms." Evolution and Human Behaviour **19**: 3-25.

Bowling, D. L., M. Hoeschele, et al. (2017). "The Nature and Nurture of Musical Consonance." Music Perception: An Interdisciplinary Journal **35**(1): 118-121.

Bowling, D. L. and D. Purves (2015). "A biological rationale for musical consonance." Proceedings of the National Academy of Sciences **112**(36): 11155-11160.

Bowling, D. L., J. Sundararajan, et al. (2012). "Expression of emotion in Eastern and Western music mirrors vocalization." PloS one **7**(3): e31942.

Boyd, B. (2018). The evolution of stories: from mimesis to language, from fact to fiction. Wiley Interdisciplinary Reviews: Cognitive Science, 9(1), e1444.

Boyd, R. and J. B. Silk (2014). How humans evolved, WW Norton & Company.

Bradbury, J. and S. Vehrencamp (1998). Principles of animal communication. Sunderland, MA, Sinauer.

Bradshaw, E. and M. A. McHenry (2005). "Pitch discrimination and pitch matching abilities of adults who sing inaccurately." Journal of Voice **19**(3): 431-439.

Brandner, H. (2012). Bewegungslinien der Musik. Alexander Truslit seine Lehre der Korpermusikalitat der Kinasthesie der Musik. Forum Musikpadogogik.

Bråten, S. (2007). On being moved: From mirror neurons to empathy, John Benjamins Publishing.

Bratman, M. (1999). Faces of intention: Selected essays on intention and agency, Cambridge University Press.

Bratman, M. E. (1989). "Intention and personal policies." Philosophical perspectives **3**: 443-469.

Bratman, M. E. (1990). "What is intention." In P. Cohen (Ed.) Intentions in communication: 15-32. MIT Press.

Bratman, M. E. (1992). "Shared cooperative activity." The philosophical review **101**(2): 327-341.

Bregman, A. (1990). Auditory scene analysis: The perceptual organization of sound. Cambridge, Massachusetts, MIT Press.

Bregman, A. S. (1994). Auditory scene analysis: The perceptual organization of sound, MIT press.

Bregman, M. R., A. D. Patel, et al. (2012). "Stimulus-dependent flexibility in non-human auditory pitch processing." Cognition **122**(1): 51-60.

Bregman, M. R., A. D. Patel, et al. (2016). "Songbirds use spectral shape, not pitch, for sound pattern recognition." Proceedings of the National Academy of Sciences **113**(6): 1666-1671.

Bretherton, I. and E. Bates (1979). "The emergence of intentional communication." New Directions for Child and Adolescent Development **1979**(4): 81-100.

Bretherton, I., S. McNew, et al. (1981). "Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a “theory of mind”." Infant social cognition: 333-373.

Brooks, D. I. and R. G. Cook (2010). "Chord discrimination by pigeons." Music Perception: An Interdisciplinary Journal **27**(3): 183-196.

Brown, C. and P. Waser (1984). "Hearing and communication in blue monkeys (Cercopithecus mitis)." Animal Behaviour **32**(11): 66-75.

Brown, D. E. (1991). Human universals, McGraw-Hill New York.

Brown, S. (2000). The "musilanguage" model of music evolution. The origins of music. N. Wallin, B. Merker and S. Brown. Cambridge Massachusetts, MIT Press**:** 271-300.

Brown, S. (2000b). Evolutionary models of music: From sexual selection to group selection. Perspectives in ethology 13: Behavior, evolution and culture. F. Tonneau and N. Thompson. New York, Plenum Publishers**:** 231-281.

Brown, S. (2001). "Are music and language homologues?" Annals-new york Academy of sciences **930**: 372-374.

Brown, S. (2007). "Contagious heterophony: a new theory about the origins of music." Musicae Scientiae **11**(1): 3-26.

Brown, S. (2017). "A joint prosodic origin of language and music." Frontiers in psychology **8**: 1894.

Brown, S. and J. Jordania (2013). "Universals in the world’s musics." Psychology of Music **41**(2): 229-248.

Brown, S., L. Martinez, et al. (2004). "Passive music listening spontaneously engages limbic and paralimbic systems." Neuroreport **15**: 2033-2037.

Brown, S. and M. J. Martinez (2007). "Activation of premotor vocal areas during musical discrimination." Brain and cognition **63**(1): 59-69.

Brown, S., M. J. Martinez, et al. (2004). "The song system of the human brain." Cognitive Brain Research **20**(3): 363-375.

Brown, S., M. J. Martinez, et al. (2006). "Music and language side by side in the brain: a PET study of the generation of melodies and sentences." European Journal of Neuroscience **23**(10): 2791-2803.

Brown, S., E. Ngan, et al. (2007). "A larynx area in the human motor cortex." Cerebral Cortex **18**(4): 837-845.

Brown, S. and U. Volgsten (2005). Music and manipulation: On the social uses and social control of music, Berghahn Books.

Brudzynski, S. M. (2009). Handbook of mammalian vocalization: An integrative neuroscience approach, Academic Press.

Bruner, J. (2003). "The narrative construction of reality." Narrative intelligence: 41-62.

Bruner, J. S. (1975). "The ontogenesis of speech acts." Journal of child language **2**(1): 1-19.

Bruner, J. S. (2009). Actual minds, possible worlds, Harvard University Press.

Brunswick, E. (1956). Perception and the representation design of psychological experiments. Berkeley, CA, University of California Press.

Bryant, G. A. and C. A. Aktipis (2014). "The animal nature of spontaneous human laughter." Evolution and Human Behavior **35**(4): 327-335.

Bryne, R. (1995). The thinking ape: Evolutionary origins of intelligence. Oxford, Oxford University Press.

Budil, I. (1994). A functional reconstruction of the supralaryngeal tract of the fossil hominid from Petralona. Studies in language origins vol. 3. J. Wind, A. Jonker, R. Allot and L. Rolfe. Amsterdam, Benjamins**:** 1-19.

Buhusi, C. V. and W. H. Meck (2005). "What makes us tick? Functional and neural mechanisms of interval timing." Nature reviews. Neuroscience **6**(10): 755.

Bulbulia, J. (2004). "The cognitive and evolutionary psychology of religion." Biology and Philosophy **19**: 655-686.

Bulbulia, J. (2004). "Religious costs as adaptations that signal altruistic intention." Evolution and Cognition **10**: 19-42.

Bulbulia, J. A., D. Xygalatas, et al. (2013). "Images from a jointly-arousing collective ritual reveal affective polarization." Frontiers in psychology **4**: 960.

Burgoon, J. K., L. K. Guerrero, et al. (2016). Nonverbal communication, Routledge.

Burnett, T. A., M. B. Freedland, et al. (1998). "Voice F0 responses to manipulations in pitch feedback." The Journal of the Acoustical Society of America **103**(6): 3153-3161.

Burnett, T. A., E. K. Hindmarsh, et al. (2007). Effect of anticipation on voice F0 responses to auditory feedback pitch shifts. Proceedings of Meetings on Acoustics 154ASA, ASA.

Burnett, T. A., J. E. Senner, et al. (1997). "Voice F 0 responses to pitch-shifted auditory feedback: a preliminary study." Journal of Voice **11**(2): 202-211.

Burns, E. and W. Ward (1999). Intervals, scales, and tuning. The psychology of music (revised edition). D. Deutsch. San Diego, Academic Press**:** 215-264.

Burns, E. M. and N. F. Viemeister (1976). "Nonspectral pitch." The Journal of the Acoustical Society of America **60**(4): 863-869.

Burns, E. M. and W. D. Ward (1999). "Intervals, scales, and tuning." The psychology of music **2**: 215-264.

Butler, J. W. and P. G. Daston (1968). "Musical consonance as musical preference: A cross-cultural study." The Journal of general psychology **79**(1): 129-142.

Buttelmann, D., F. Buttelmann, et al. (2017). "Great apes distinguish true from false beliefs in an interactive helping task." PloS one **12**(4): e0173793.

Buttelmann, D., J. Call, et al. (2009). "Do great apes use emotional expressions to infer desires?" Developmental science **12**(5): 688-698.

Buttelmann, D., M. Carpenter, et al. (2007). "Enculturated chimpanzees imitate rationally." Developmental science **10**(4).

Buttelmann, D., M. Carpenter, et al. (2008). "Rational tool use and tool choice in human infants and great apes." Child development **79**(3): 609-626.

Buttelmann, D., M. Carpenter, et al. (2009). "Eighteen-month-old infants show false belief understanding in an active helping paradigm." Cognition **112**(2): 337-342.

Buttelmann, D., S. Schütte, et al. (2012). "Great apes infer others' goals based on context." Animal cognition **15**(6): 1037-1053.

Call, J., B. Hare, et al. (2004). "'Unwilling'versus 'unable': chimpanzees' understanding of human intentional action." Developmental science **7**(4): 488-498.

Call, J., B. A. Hare, et al. (1998). "Chimpanzee gaze following in an object-choice task." Animal cognition **1**(2): 89-99.

Call, J. and M. Tomasello (2008). "Does the chimpanzee have a theory of mind? 30 years later." Trends in cognitive sciences **12**(5): 187-192.

Calvo-Merino, B., J. Grèzes, et al. (2006). "Seeing or doing? Influence of visual and motor familiarity in action observation." Current Biology **16**(19): 1905-1910.

Cameron, D. J. and J. A. Grahn (2016). "The Neuroscience of Rhythm." The Oxford handbook of music psychology: 357.

Cann, R., M. Stoneking, et al. (1987). "Mitochondrial DNA and human evolution." Nature **325**: 31-36.

Capps, M. J. and H. W. Ades (1968). "Auditory frequency discrimination after transection of the olivocochlear bundle in squirrel monkeys." Experimental neurology **21**(2): 147-158.

Capranica, R. (1977). Auditory processing of vocal signals in anurans. Reproductive biology of amphibians. D. Taylor and S.Guttman. New York, Plenum**:** 303-317.

Caron, A. J., E. J. Kiel, et al. (2002). "Comprehension of the referential intent of looking and pointing between 12 and 15 months." Journal of Cognition and Development **3**(4): 445-464.

Carpenter, M., N. Akhtar, et al. (1998). "Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions." Infant behavior and development **21**(2): 315-330.

Carpenter, M. and M. Tomasello (1995). "Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees." Social Development **4**(3): 217-237.

Carpenter, M., M. Tomasello, et al. (2005). "Role reversal imitation and language in typically developing infants and children with autism." Infancy **8**(3): 253-278.

Carter, C. S. (2005). Attachment and bonding: A new synthesis, Mit Press.

Carter, R. (1998). Mapping the mind. London, Weidenfeld & Nicholson.

Castellano, M., J. Bharucha, et al. (1984). "Tonal hierarchies in the music of North India." Journal of Experimental Psychology: General **113**: 394-412.

Catchpole, C. and P. Slater (1995). Bird song: Biological themes and variations. Cambridge, Cambridge University Press.

Catchpole, C. K. and P. J. Slater (2003). Bird song: biological themes and variations, Cambridge university press.

Cavalli-Sforza, L. (1998). "The DNA revolution in population genetics." Trends in Genetics **14**: 60-65.

Cavalli-Sforza, L. and M. Feldman (2003). "The application of molecular genetic approaches to the study of human evolution." Nature Genetics Supplement **33**: 266-275.

Chang, H. and S. Trehub (1977a). "Auditory processing of relational information by young infants." Journal of Experimental Child Psychology **24**: 324-331.

Chang, H. and S. Trehub (1977b). "Infants' perception of temporal grouping in auditory patterns." Child Develoment **48**: 1666-1670.

Chartrand, T. L. and J. L. Lakin (2013). "The antecedents and consequences of human behavioral mimicry." Annual review of psychology **64**: 285-308.

Chase, A. (2001). "Music discriminations by carp (Cyprinus carpio)." Animal Learning and Behavior **29**(4): 336-353.

Chase, P. and A. Novell (1998). "Taphonomony of a suggested Middle Paleolithic bone flute from Slovenia." Current Anthropology **39**(4): 184-194.

Chen, J. L., V. B. Penhune, et al. (2008). "Listening to musical rhythms recruits motor regions of the brain." Cerebral cortex **18**(12): 2844-2854.

Chen, J. L., V. B. Penhune, et al. (2008). "Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training." Journal of cognitive neuroscience **20**(2): 226-239.

Chen, J. L., R. J. Zatorre, et al. (2006). "Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms." Neuroimage **32**(4): 1771-1781.

Chen, Q., J. B. Heston, et al. (2013). "Expression analysis of the speech-related genes FoxP1 and FoxP2 and their relation to singing behavior in two songbird species." Journal of Experimental Biology **216**(19): 3682-3692.

Chen, S. H., H. Liu, et al. (2007). "Voice F 0 responses to pitch-shifted voice feedback during English speech." The Journal of the Acoustical Society of America **121**(2): 1157-1163.

Cheney, D. L. and R. M. Seyfarth (1985). "Vervet monkey alarm calls: manipulation through shared information?" Behaviour **94**(1): 150-166.

Chiandetti, C. and G. Vallortigara (2011). "Chicks like consonant music." Psychological science **22**(10): 1270-1273.

Chomsky, N. (1993). Language and thought. London, Moyer Bell.

Chomsky, N. (2014). Aspects of the Theory of Syntax, MIT press.

Chomsky, N., M. Hauser, et al. (2002). "The faculty of language: What is it? Who has it? And how did it evolve?" Science **298**(5598): 1569-1669.

Chow, C. P., J. F. Mitchell, et al. (2015). Vocal turn-taking in a non-human primate is learned during ontogeny. Proc. R. Soc. B, The Royal Society.

Christensen, T. (2004). Rameau and musical thought in the Enlightenment, Cambridge University Press.

Cirelli, L. K., K. M. Einarson, et al. (2014). "Interpersonal synchrony increases prosocial behavior in infants." Developmental science **17**(6): 1003-1011.

Cirelli, L. K., S. J. Wan, et al. (2017). "Effects of Interpersonal Movement Synchrony on Infant Helping Behaviors: Is Music Necessary?" Music Perception: An Interdisciplinary Journal **34**(3): 319-326.

Cirelli, L. K., S. J. Wan, et al. (2014). "Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness." Phil. Trans. R. Soc. B **369**(1658): 20130400.

Cirelli, L. K., S. J. Wan, et al. (2016). "Social effects of movement synchrony: increased infant helpfulness only transfers to affiliates of synchronously moving partners." Infancy **21**(6): 807-821.

Clark, A., S. Glanowski, et al. (2003). "Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios." Science **302**(5652): 1960-1963.

Clark, C., H. Golden, et al. (2015). "Acquired amusia." Handbook of clinical neurology **129**: 607-631.

Clarkson, M. and R. Clifton (1995). "Infants' pitch perception: inharmonic tonal complexes." Journal of the Acoustical Society of America **98**: 1372-1379.

Clarkson, M. G. and R. K. Clifton (1985). "Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental." The Journal of the Acoustical Society of America **77**(4): 1521-1528.

Clarkson, M. G. and E. C. Rogers (1995). "Infants require low-frequency energy to hear the pitch of the missing fundamental." The Journal of the Acoustical Society of America **98**(1): 148-154.

Clay, Z., J. Archbold, et al. (2015). "Functional flexibility in wild bonobo vocal behaviour." PeerJ **3**: e1124.

Clayton, M. (2012). "What is entrainment? Definition and applications in musical research." Empirical musicology review. **7**(1-2): 49-56.

Clayton, M. (2016). "The social and personal functions of music in cross-cultural perspective." The Oxford handbook of music psychology: 35-44.

Clayton, M., T. Herbert, et al., Eds. (2003). The cultural study of music. Oxford, Routledge.

Clayton, M., R. Sager, et al. (2004). "In time with the music: The concept of entrainment and its significance for ethnomusicology." ESEM counterpoint **1**: 1-45.

Clayton, M. R. (1996). "Free rhythm: ethnomusicology and the study of music without metre." Bulletin of the School of Oriental and African Studies **59**(2): 323-332.

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

Clynes, M. (1977). Sentics: The touch of emotions. New York, Anchor Press/Doubleday.

Clynes, M. and J. Walker (1982). Neurobiological functions of rhythm, time, and pulse in music. Music, mind, and brain: The neuropsychology of music. M. Clynes. New York, Plenum.

Cochrane, T., B. Fantini, et al. (2013). The emotional power of music: Multidisciplinary perspectives on musical arousal, expression, and social control, OUP Oxford.

Coey, C., M. Varlet, et al. (2011). "Effects of movement stability and congruency on the emergence of spontaneous interpersonal coordination." Experimental brain research **211**(3-4): 483-493.

Coffey, E. B., S. C. Herholz, et al. (2016). "Cortical contributions to the auditory frequency-following response revealed by MEG." Nature communications **7**.

Cohen, A. (2001). Music as a source of emotion in film. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press**:** 249-273.

Cohen, A., L. Thorpe, et al. (1987). "Infants' perception of musical relations in short transposed tone sequences." Canadian Journal of Psychology **41**: 33-47.

Cole, J. and E. Sedgwick (1992). "The perceptions of force and of movement in a man without large myelinated sensory afferents below the neck." The Journal of physiology **449**(1): 503-515.

Collier, G. L. (1996). "Microrhythms in jazz: a review of papers." Annual review of jazz studies: 117-139.

Collier, G. L. and J. L. Collier (2002). "A study of timing in two Louis Armstrong solos." Music Perception: An Interdisciplinary Journal **19**(3): 463-483.

Coltheart, M. (1999). "Modularity and cognition." Trends in cognitive sciences **3**(3): 115-120.

Conard, N. J., M. Malina, et al. (2009). "New flutes document the earliest musical tradition in southWestern Germany." Nature **460**(7256): 737-740.

Connolly, C. and A. Williamon (2004). "Mental skills training." Musical excellence: Strategies and techniques to enhance performance **5**(1): 221-245.

Connor, R. C., R. Smolker, et al. (2006). "Synchrony, social behaviour and alliance affiliation in Indian Ocean bottlenose dolphins, Tursiops aduncus." Animal Behaviour **72**(6): 1371-1378.

Cook, P., A. Rouse, et al. (2013). "A California sea lion (Zalophus californianus) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic." Journal of Comparative Psychology **127**(4): 412.

Cooke, D. (1959). The language of music. Oxford, Oxford University Press.

Cooke, J., S. Brown, et al. (1985). "Initial agonist burst duration changes with movement amplitude in a deafferented patient." Experimental brain research **60**(1): 184-187.

Coolidge, F. L. and T. Wynn (2005). "Working memory, its executive functions, and the emergence of modern thinking." Cambridge archaeological journal **15**(1): 5-26.

Coolidge, F. L., T. Wynn, et al. (2013). "The evolution of working memory." Working memory: the connected intelligence: 37-59.

Corballis, M. C. (2014). The recursive mind: The origins of human language, thought, and civilization, Princeton University Press.

Corballis, M. C. (2017). "The evolution of language: sharing our mental lives." Journal of Neurolinguistics **43**: 120-132.

Corrigall, K. A. and L. J. Trainor (2010). "Musical enculturation in preschool children: Acquisition of key and harmonic knowledge." Music Perception: An Interdisciplinary Journal **28**(2): 195-200.

Costa-Giomi, E. (2004). "Effects of three years of piano instruction on children’s academic achievement, school performance and self-esteem." Psychology of Music **32**(2): 139-152.

Costa-Giomi, E. (2003). "Young children's harmonic perception." Annals of the New York Academy of Sciences **999**(1): 477-484.

Coudé, G., P. F. Ferrari, et al. (2011). "Neurons controlling voluntary vocalization in the macaque ventral premotor cortex." PloS one **6**(11): e26822.

Coull, J. and A. Nobre (2008). "Dissociating explicit timing from temporal expectation with fMRI." Current opinion in neurobiology **18**(2): 137-144.

Cousineau, M., J. H. McDermott, et al. (2012). "The basis of musical consonance as revealed by congenital amusia." Proceedings of the National Academy of Sciences **109**(48): 19858-19863.

Cousineau, M., A. J. Oxenham, et al. (2015). "Congenital amusia: a cognitive disorder limited to resolved harmonics and with no peripheral basis." Neuropsychologia **66**: 293-301.

Cramer, E. M. and W. Huggins (1958). "Creation of pitch through binaural interaction." The Journal of the Acoustical Society of America **30**(5): 413-417.

Cravo, A. M., G. Rohenkohl, et al. (2013). "Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex." Journal of Neuroscience **33**(9): 4002-4010.

Crespo-Bojorque, P. and J. M. Toro (2015). "The use of interval ratios in consonance perception by rats (Rattus norvegicus) and humans (*Homo sapiens*)." Journal of Comparative Psychology **129**(1): 42.

Crespo-Bojorque, P. and J. M. Toro (2016). "Processing advantages for consonance: A comparison between rats (Rattus norvegicus) and humans (*Homo sapiens*)." Journal of Comparative Psychology **130**(2): 97.

Crockford, C., R. M. Wittig, et al. (2012). "Wild chimpanzees inform ignorant group members of danger." Current Biology **22**(2): 142-146.

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

Cross, I. (2001). "Music, cognition, culture, and evolution." Annals of the New York Academy of sciences **930**(1): 28-42.

Cross, I. (2003a). Music, cognition, culture and evolution. The cognitive science of music. I. Peretz and R. Zatorre. Oxford, Oxford University Press**:** 42-56.

Cross, I. (2003b). Music and biocultural evolution. The cultural study of music: A critical introduction. M. Clayton, T. Herbert and R. Middleton. London, Routledge**:** 19-30.

Cross, I. (2005). Music and meaning, ambiguity and evolution. Musical communication. D. Miell, R. MacDonald and D. Hargreaves. Oxford, Oxford University Press.

Cross, I. (2007). "Music and cognitive evolution." Oxford handbook of evolutionary psychology: 649-667.

Cross, I. (2008). "Musicality and the human capacity for culture." Musicae Scientiae **12**(1_suppl): 147-167.

Cross, I. (2009). "The evolutionary nature of musical meaning." Musicae Scientiae **13**(2_suppl): 179-200.

Cross, I. (2010). "The evolutionary basis of meaning in music: some neurological and neuroscientific implications." The neurology of music: 1-15.

Cross, I. (2016). "The nature of music and its evolution." The Oxford handbook of music psychology (2nd Edition): 3-13.

Cross, I., J. Bispham, et al. (2008). "The Evolution of Musical Rhythm." Submitted to Evolutionary Anthropology (Unpublished) - https://cambridge.academia.edu/JohnBispham.

Cross, I., T. Fitch, et al. (2013). Culture and evolution, MIT Press. In: Arbib, M, (ed.) Language, Music and the Brain: A mysterious relationship.

Cross, I. and D. Kronhaus (2007). "Global musical structure and the attribution of affect." Unpublished Manuscript.

Cross, I. and I. Morley (2009). "The evolution of music: Theories, definitions and the nature of the evidence." Communicative musicality: Exploring the basis of human companionship: 61-81.

Cross, I. and M. Rohrmeier, Eds. (2007). Proceedings of the SEMPRE conference on Music and Language. University of Cambridge, UK.

Cross, I., E. Zubrow, et al. (2002). Musical behaviours and the archaeological record: a preliminary study. Experimental Archaeology. British archaeological reports international series. J. Mathieu. **1035:** 25-34.

Crowder, R. (1993). Auditory Memory: In 'Thinking in sound: The cognitive psychology of human audition'. Based on the fourth workshop in the Tutorial Workshop series organized by the Hearing Group of the French Acoustical Society., Clarendon Press/Oxford University Press.

Csibra, G., S. Bıró, et al. (2003). "One-year-old infants use teleological representations of actions productively." Cognitive science **27**(1): 111-133.

Csibra, G., G. Gergely, et al. (1999). "Goal attribution without agency cues: the perception of ‘pure reason’in infancy." Cognition **72**(3): 237-267.

Csikszentmihalyi, M. (1990). Flow. The psychology of optimal experience. New York, Harper & Row.

Csikszentmihalyi, M. and K. Rathunde (1993). "The measurement of flow in everyday life: toward a theory of emergent motivation."

Cuddy, L. and B. Badertscher (1987). "Recovery of the tonal hierarchy: Some comparisions across age and levels of musical experience." Perception and Psychophysics **41**: 609-620.

Curtis, M. E. and J. J. Bharucha (2010). "The minor third communicates sadness in speech, mirroring its use in music." Emotion **10**(3): 335.

Cusick, S. G. (2006). Music as torture, music as weapon. http://www.sibetrans.com/trans/article/152/music-as-torture-music-as-weapon.

Cynx, J. (1993). "Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (Sturnus vulgaris)." Journal of Comparative Psychology **107**(2): 140.

Cynx, J. and M. Shapiro (1986). "Perception of missing fundamental by a species of songbird (Sturnus vulgaris)." J Comp Psychol **100**(4): 356-360.

D'ausilio, A., E. Altenmüller, et al. (2006). "Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece." European Journal of Neuroscience **24**(3): 955-958.

D'Errico, F. (1998). No holes in my argument. The contribution of microscopy and taphonomy to the interpretation of the Divje Babe "flute". The question of the mousterian bone flute from Divje babe I cave. Spodnja Idrija Congress. Ljubijana, Slovenia.

D'Errico, F., C. Henschilwood, et al. (2003). "Archaeological evidence for the emergence of language, symbolism and music - an alternative multidisciplinary perspective." Journal of Prehistory **17**: 1-70.

D'Errico, F., P. Villa, et al. (1998). "A Middle Paleolithic origin of music? Using cave bear bone accumulations to assess the Divje Babe I cave bone "flute"." Antiquity **72**: 65-79.

D'esposito, M. and B. R. Postle (2015). "The cognitive neuroscience of working memory." Annual review of psychology **66**.

Dainton, B. (2013). "The perception of time." A companion to the philosophy of time: 387-409.

Dalla Bella, S., J.-F. Giguère, et al. (2007). "Singing proficiency in the general population." The Journal of the Acoustical Society of America **121**(2): 1182-1189.

Dalla Bella, S., J.-F. Giguère, et al. (2009). "Singing in congenital amusia." The Journal of the Acoustical Society of America **126**(1): 414-424.

Dalla Bella, S., I. Peretz, et al. (2001). "A developmental study of the affective value of tempo and mode in music." Cognition **80**(3): B1-B10.

Damasio, A. R. (2000). The Feeling of what Happens: Body, Emotion and the Making of Consciousness, Random House.

Dams, L. (1984). "Preliminary findings at the "organ sanctuary" in the cave of Nerja, Malaga, Spain." Oxford Journal of Archaeology **3**: 1-14.

Dams, L. (1985). "Paleolithic lithophones: Descriptions and comparisons." Oxford Journal of Archaeology: 31-46.

Danesi, M. and P. Perron (1999). Analyzing cultures: An introduction and handbook, Indiana University Press.

Darwin, C. (1859). On the Origin of Species by Natural Selection. London, Murray.

Darwin, C. (1871). The descent of man and selection in relation to sex. London, John Murray.

Darwin, C. (1872). The expression of emotions of man and animals. London, Harper Collins.

Darwin, C. (1872). The expression of the emotions in man and animals John Murray, London.

Darwin, C. (2005). "Pitch and auditory grouping." Pitch: 278-305.

Das, M., D. M. Howard, et al. (1999). Motion curves in music: the statistical analysis of MIDI data. EUROMICRO Conference, 1999. Proceedings. 25th, IEEE.

Dauer, R. M. (1983). "Stress-timing and syllable-timing reanalyzed." Journal of phonetics.

Dauvois, M. (1989). "Son et musique paleolithiques." Les Dossiers D'Archeologie **142**: 2-11.

David, A. (1967). Elements of general phonetics, Edinburgh: Edinburgh University Press.

Davidson, J. (1994). On emotion, mood, and related affective constructs. The nature of emotion: Fundamental questions. P.

Ekman and R. Davidson. New York, Oxford University Press**:** 51-55.

Davidson, R., K. Scherer, et al., Eds. (2003). Handbook of affective sciences. Oxford, Oxford University Press.

Davidson, R. J., D. A. Lewis, et al. (2002). "Neural and behavioral substrates of mood and mood regulation." Biological psychiatry **52**(6): 478-502.

Davidson, R. J., K. R. Sherer, et al. (2009). Handbook of affective sciences, Oxford University Press.

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

Davies, M., G. Madison, et al. (2013). "The effect of microtiming deviations on the perception of groove in short rhythms." Music Perception: An Interdisciplinary Journal **30**(5): 497-510.

Davies, S. (2010). "Emotions Expressed and Aroused by Music: Philosophical Perspectives." Handbook of Music and Emotion: Theory, Research, Applications.

Davis, W. and M. Thaut (1989). "The influence of preferred relaxing music on measures of state anxiety, relaxation, and physiological measures." Journal of Music Therapy **26**: 168-187.

De Cheveigne, A. (2005). Pitch perception models. Pitch, Springer**:** 169-233.

de Cheveigné, A. (2010). "Pitch perception." The oxford handbook of auditory science: Hearing: 71-104.

De Dreu, M., A. Van Der Wilk, et al. (2012). "Rehabilitation, exercise therapy and music in patients with Parkinson's disease: a meta-analysis of the effects of music-based movement therapy on walking ability, balance and quality of life." Parkinsonism & related disorders **18**: S114-S119.

De Jaegher, H. and E. Di Paolo (2007). "Participatory sense-making." Phenomenology and the cognitive sciences **6**(4): 485-507.

De Manzano, Ö., T. Theorell, et al. (2010). "The psychophysiology of flow during piano playing." Emotion **10**(3): 301.

de Vos, C., F. Torreira, et al. (2015). "Turn-timing in signed conversations: coordinating stroke-to-stroke turn boundaries." Frontiers in psychology **6**: 628.

de Vries, B. (1991). "Assessment of the affective response to music with Clynes's sentograph." Psychology of Music **19**: 46-64.

De Waal, F. B. (1988). "The communicative repertoire of captive bonobos (Pan paniscus), compared to that of chimpanzees." Behaviour **106**(3): 183-251.

Deacon, T. (1997). The symbolic species. London, Allen Lane.

Deacon, T. (1997). The Symbolic Species: The co-evolution of language and the human brain. 527 pp, The Penguin Press: London.

Dean, R. T., F. Bailes, et al. (2008). "Microtonality, the octave, and novel tunings for affective music." Music of the Spirit: Asian-Pacific Musical Identity: 128-139.

Decety, J. and J. Grèzes (2006). "The power of simulation: imagining one's own and other's behavior." Brain research **1079**(1): 4-14.

Decety, J. and J. A. Sommerville (2003). "Shared representations between self and other: a social cognitive neuroscience view." Trends in cognitive sciences **7**(12): 527-533.

DeGusta, D., W. Gilbert, et al. (1999). "Hypoglossal canal size and hominid speech." Proceedings of the National Academy of Sciences **96**: 1800-1804.

Dehaene, S. and J.-P. Changeux (2011). "Experimental and theoretical approaches to conscious processing." Neuron **70**(2): 200-227.

Demany, L. and F. Armand (1984). "The perceptual reality of tone chroma in early infancy." The Journal of the Acoustical Society of America **76**(1): 57-66.

Demany, L., B. McKenzie, et al. (1977). "Rhythm perception in early infancy." Nature **266**: 718-719.

Demolin, D. (2006). "The evolution of language and music in comparative perspective." unpublished.

Demolin, D. and V. Delvaux (2006). A comparison of the articulatory parameters involved in the production of sounds of bonobos and modern humans. The evolution of language: Proceedings of the 6th International Conference (Evolang6).

Demos, A. P., R. Chaffin, et al. (2012). "Rocking to the beat: Effects of music and partner's movements on spontaneous interpersonal coordination." Journal of Experimental Psychology: General **141**(1): 49.

Dennett, D. (1998). Brainchildren: Essays on designing minds. London, Penguin.

DeNora, T. (2001). Aesthetic agency and musical practice: New directions on the sociology of music and emotion. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press.

Deutsch, D. (Ed.) (2013). The Psychology of Music (3rd Edition). Academic Press.

Deutsch, D., T. Henthorn, et al. (2011). "Illusory transformation from speech to song a." The Journal of the Acoustical Society of America **129**(4): 2245-2252.

Dewey, J. A. and G. Knoblich (2016). "17 Representation of Self versus Others' Actions." Shared Representations: Sensorimotor Foundations of Social Life: 351.

Di Pietro, M., M. Laganaro, et al. (2004). "Receptive amusia: Temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion." Neuropsychologia **42**: 868-877.

Diedrich, C. G. (2015). "'Neanderthal bone flutes': simply products of Ice Age spotted hyena scavenging activities on cave bear cubs in European cave bear dens." Open Science **2**(4): 140022.

Diedrichsen, J., R. B. Ivry, et al. (2003). "Cerebellar and basal ganglia contributions to interval timing." Functional and neural mechanisms of interval timing: 457-481.

Disotell, T. (1999). "Origins of modern humans still look recent." Current Biology **9**(647-650).

Dissanayake, E. (1979). "An ethological view of ritual and art in human evolutionary history." Leonardo **12**(1): 27-31.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interaction. The origins of music. B. M. N. Wallin, & S. Brown. MA, MIT Press.

Dissanayake, E. (2000). Art and intimacy: How the arts began, University of Washington Press.

Dissanayake, E. (2001). Homo aestheticus: Where art comes from and why, University of Washington Press.

Dissanayake, E. (2008). "If music is the food of love, what about survival and reproductive success?" Musicae Scientiae **12**(1_suppl): 169-195.

Dissanayake, E. (2015). What is art for?, University of Washington Press.

Dodenhoff, D. J., R. D. Stark, et al. (2001). "Do woodpecker drums encode information for species recognition?" The Condor **103**(1): 143-150.

Doffman, M. (2013). Groove: temporality, awareness and the feeling of entrainment in jazz performance. Experience and Meaning in Music Performance, Oxford University Press**:** 62-85.

Donald, M. (1991). Origins of the modern mind: three stages in the evolution of culture and cognition. Cambridge, Mass.: Harvard University Press.

Donald, M. (2017). Key cognitive preconditions for the evolution of language. Psychonomic bulletin & review, 24(1), 204-208.

Donath, T. M., U. Natke, et al. (2002). "Effects of frequency-shifted auditory feedback on voice F 0 contours in syllables." The Journal of the Acoustical Society of America **111**(1): 357-366.

Donnet, S., R. Bartolo, et al. (2014). "Monkeys time their pauses of movement and not their movement-kinematics during a synchronization-continuation rhythmic task." Journal of neurophysiology **111**(10): 2138-2149.

Dor, D. (2015). The instruction of imagination: Language as a social communication technology, Foundations of Human Interacti.

Dorit, R., H. Akashi, et al. (1995). "Absence of polymorphism at the ZFY locus on the human Y chromosome." Science **268**(1183-1185).

Doumas, M. and A. M. Wing (2007). "Timing and trajectory in rhythm production." Journal of experimental psychology: Human perception and performance **33**(2): 442.

Doupe, A. J., D. J. Perkel, et al. (2005). "Birdbrains could teach basal ganglia research a new song." Trends in neurosciences **28**(7): 353-363.

Dowling, W. and D. Harwood (1986). Music cognition. London, Academic Press.

Dowling, W. and D. Harwood (1986). Music cognition. Series in cognition and perception, Academic Press.

Drake, C. and D. Bertrand (2001). "The quest for universals in temporal processing in music." Annals of the New York Academy of sciences **930**(1): 17-27.

Drayna, D., A. Manichaikul, et al. (2001). "Genetic correlates of musical pitch recognition in humans." Science **291**(5510): 1969-1972.

Duchin, L. (1990). "The evolution of articulate speech: comparative anatomy of Pan and Homo." Journal of Human Evolution **19**: 687-697.

Dugatkin, L. and H. Reeve (1994). "Behavioural ecology and the levels of selection: Dissolving the group selection controversy." Advances in the Study of Behavior **23**: 101-133.

Dujardin, E. and U. Jürgens (2005). "Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey." Brain research **1034**(1): 114-131.

Dunbar, R. (1997). Grooming, gossip and the evolution of language. New York, Faber & Faber.

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

Dunbar, R. (2014). Mind the gap: or why humans aren't just great apes. Lucy to language: The benchmark papers, Oxford University Press**:** 3-18.

Dunbar, R., C. Knight, et al., Eds. (1999). The evolution of culture. Edinburgh, Edinburgh University Press.

Dunbar, R. I. (1991). "Functional significance of social grooming in primates." Folia primatologica **57**(3): 121-131.

Dunbar, R. I. (2003). "The social brain: mind, language, and society in evolutionary perspective." Annual Review of Anthropology **32**(1): 163-181.

Dunbar, R. I. (2010). "The social role of touch in humans and primates: behavioural function and neurobiological mechanisms." Neuroscience & Biobehavioral Reviews **34**(2): 260-268.

Dunlap, K. (1910). "Reaction to rhythmic stimuli with attempt to synchronize." Psychological review **17**(6): 399.

Durham, W. (1991). Coevolution: Genes, culture and human diversity. Stanford, Stanford University Press.

Durkheim, E. (1912). (1995). The Elementary Forms of the Religious Life, trans. Karen Fields, New York: Free Press.

Durkheim, E. (1912/1926). The elementary forms of the religious life: A study in religious sociology (J. Swain, Trans.). New York, Macmillan.

Eerola, T., T. Himberg, et al. (2006). "Perceived complexity of Western and African folk melodies by Western and African listeners." Psychology of Music **34**(3): 337-371.

Eerola, T. and J. K. Vuoskoski (2013). "A review of music and emotion studies: approaches, emotion models, and stimuli." Music Perception: An Interdisciplinary Journal **30**(3): 307-340.

Egnor, S. R. and M. D. Hauser (2004). "A paradox in the evolution of primate vocal learning." Trends in neurosciences **27**(11): 649-654.

Ehrenreich, B. (2007). Dancing in the streets: A history of collective joy, Metropolitan books.

Einarson, K. M. (2017). Beat perception and synchronization abilities in young children.

Eisler, H., A. D. Eisler, et al. (2008). "Psychophysical issues in the study of time perception." Psychology of time: 75-109.

Ekman, P. (1992). "An argument for basic emotions." Cognition and Emotion **6**: 169-200.

Ekman, P. and J. Davidson, Eds. (1994). The nature of emotion: Fundamental questions. New York, Oxford University Press.

Elie, J. E. and F. E. Theunissen (2016). "The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals." Animal cognition **19**(2): 285-315.

Elliot, A. J. and M. V. Covington (2001). "Approach and avoidance motivation." Educational Psychology Review **13**(2): 73-92.

Ellis, C. (1985). Aboriginal music: Education for living. St. Lucia, University of Queensland Press.

Ellis, R. J. and J. F. Thayer (2010). "Music and autonomic nervous system (dys) function." Music Perception: An Interdisciplinary Journal **27**(4): 317-326.

Elman, J., E. Bates, et al. (1996). Rethinking innateness: A connectionist perspective on development. Cambridge, MA, MIT Press.

Ember, C. (1978). "Myths about hunter-gatherers." Ethnology **17**: 439-448.

Enquist, M. (1985). "Communication during aggressive interactions with particular reference to variation in choice of behaviour." Animal Behaviour **33**: 1152-1161.

Epstein, D. (1994). Shaping time: Music, the brain, and performance. New York, Schirmer.

Estis, J. M., J. K. Coblentz, et al. (2009). "Effects of increasing time delays on pitch-matching accuracy in trained singers and untrained individuals." Journal of Voice **23**(4): 439-445.

Estis, J. M., A. Dean-Claytor, et al. (2011). "Pitch-matching accuracy in trained singers and untrained individuals: the impact of musical interference and noise." Journal of Voice **25**(2): 173-180.

Everett, D., B. Berlin, et al. (2005). "Cultural constraints on grammar and cognition in Piraha: Another look at the design features of human language." Current Anthropology **46**(4): 621-646.

Everett, D. L. (2012). Language: The cultural tool, Vintage.

Fabian, D., R. Timmers, et al. (2014). Expressiveness in music performance: Empirical approaches across styles and cultures, Oxford University Press, USA.

Fadiga, L., L. Craighero, et al. (2002). "Speech listening specifically modulates the excitability of tongue muscles: a TMS study." European Journal of Neuroscience **15**(2): 399-402.

Fairhurst, M. T., P. Janata, et al. (2012). "Being and feeling in sync with an adaptive virtual partner: brain mechanisms underlying dynamic cooperativity." Cerebral cortex **23**(11): 2592-2600.

Faith, M. and J. Thayer (2001). A dynamic systems model of musically induced emotions: Physiological and self-report evidence. The biological foundations of music. R. Zatorre and I. Peretz. New York, The New York Academy of Sciences. **930**.

Falk, D. (1992). Braindance: New discoveries about human origins and brain evolution. New York, Henry Holt and Co.

Falk, D. (2004). "Prelinguistic evolution in hominin mothers and babies: Whence motherese?" Behavioural and Brain Sciences **27**(4): 491-503.

Farabaugh, S. (1982). The ecological and social significance of duetting. Acoustic communication in birds. E. H. Miller and H. Ouellett. New York, Academic Press**:** 85-124.

Farah, M., K. Wilson, et al. (1998). "What is "special" about face perception?" Psychological Review **105**(3): 482-498.

Faulkner, A. (1985). "Pitch discrimination of harmonic complex signals: Residue pitch or multiple component discriminations?" The Journal of the Acoustical Society of America **78**(6): 1993-2004.

Fawcett, C. and U. Liszkowski (2012). "Mimicry and play initiation in 18-month-old infants." Infant behavior and development **35**(4): 689-696.

Fedurek, P., A. M. Schel, et al. (2013). "The acoustic structure of chimpanzee pant-hooting facilitates chorusing." Behavioral Ecology and Sociobiology **67**(11): 1781-1789.

Feld, S. (1982). Sound and sentiment: Birds, weeping, poetics and song in Kaluli expression. Philadelphia, University of Pennsylvania Press.

Feldman, J., D. Epstein, et al. (1992). "Force dynamics of tempo change in music." Music Perception: An Interdisciplinary Journal **10**(2): 185-203.

Feldman, R. (2003). "Infant–mother and infant–father synchrony: The coregulation of positive arousal." Infant Mental Health Journal **24**(1): 1-23.

Feldman, R. (2006). "From biological rhythms to social rhythms: Physiological precursors of mother-infant synchrony." Developmental Psychology **42**(1): 175.

Feldman, R. (2007). "Parent–infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions." Journal of Child psychology and Psychiatry **48**(34): 329-354.

Feldman, R. (2009). "The development of regulatory functions from birth to 5 years: Insights from premature infants." Child development **80**(2): 544-561.

Feldman, R., I. Gordon, et al. (2011). "Maternal and paternal plasma, salivary, and urinary oxytocin and parent–infant synchrony: considering stress and affiliation components of human bonding." Developmental science **14**(4): 752-761.

Feldman, R., R. Magori-Cohen, et al. (2011). "Mother and infant coordinate heart rhythms through episodes of interaction synchrony." Infant behavior and development **34**(4): 569-577.

Feldman, R., M. Singer, et al. (2010). "Touch attenuates infants' physiological reactivity to stress." Developmental science **13**(2): 271-278.

Feng, L. and X. Wang (2017). "Harmonic template neurons in primate auditory cortex underlying complex sound processing." Proceedings of the National Academy of Sciences: 201607519.

Fernald, A. (1991). "Prosody in speech to children: Prelinguistic and linguistic functions." Annals of Child Development **8**: 43-80.

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

Fernald, A., N. Kermanschachi, et al. (1984). The rhythms and sounds of soothing: Maternal vesitbular, tactile and auditory stimulation and infant state. International Conference on Infant Studies. New York.

Fernald, A. and H. Morikawa (1993). "Common themes and cultural variations in Japanese and American mothers' speech to infants." Child Development **64**(3): 637-656.

Fessler, D. M. and C. Holbrook (2016). "Synchronized behavior increases assessments of the formidability and cohesion of coalitions." Evolution and Human Behavior **37**(6): 502-509.

Field, T. (1994). "The effects of mother's physical and emotional unavailability on emotion regulation." Monographs of the society for research in child development **59**(2-3): 208-227.

Fink, R. (1997). "Neanderthal Flute: Oldest Musical Instrument's 4 Notes Matches 4 of Do, Re, Mi Scale." Retrieved from web site: http://www.greenwych. ca/fl-compl. htm.

Fischer, J. (1998). "Barbury macaques categorize shrill barks into two call types." Animal Behaviour **55**: 799-807.

Fischer, J. and K. Hammerschmidt (2011). "Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: insights into the evolution of vocal communication." Genes, Brain and Behavior **10**(1): 17-27.

Fischer, R., R. Callander, et al. (2013). "How do rituals affect cooperation?" Human Nature **24**(2): 115-125.

Fischer, R., D. Xygalatas, et al. (2014). "The fire-walker's high: Affect and physiological responses in an extreme collective ritual." PloS one **9**(2): e88355.

Fischman, J. (1996). "Evidence mounts for our African origin - and alternatives." Science **271**: 1364.

Fisher, R. (1930). The genetical theory of natural selection. Oxford, Clarendon Press.

Fishman, Y. I., C. Micheyl, et al. (2013). "Neural representation of harmonic complex tones in primary auditory cortex of the awake monkey." Journal of Neuroscience **33**(25): 10312-10323.

Fishman, Y. I., I. O. Volkov, et al. (2001). "Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans." Journal of neurophysiology **86**(6): 2761-2788.

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

Fitch, T. and D. Ruby (2001). "The descended larynx is not uniquely human." Proceedings of the Royal Society of London B **268**: 1669-1675.

Fitch, W. (2006). "Production of vocalizations in mammals." Visual Communication **3**(2006): 145.

Fitch, W. and M. Hauser (1995). "Vocal production in nonhuman primates: acoustics, physiology, and function constraints on "honest" advertisement." American Journal of Primatology **37**: 191-219.

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

Fitch, W. T. (2005). "The evolution of language: a comparative review." Biology and philosophy **20**(2-3): 193-203.

Fitch, W. T. (2011). "Speech perception: A language-trained chimpanzee weighs in." Current Biology **21**(14): R543-R546.

Fitch, W. T. (2012). "The biology and evolution of rhythm: Unraveling a paradox." Language and music as cognitive systems: 73-95.

Fitch, W. T. (2013). "Musical protolanguage: Darwin's theory of language evolution revisited." Birdsong, speech, and language: Exploring the evolution of mind and brain: 489-504.

Fitch, W. T. (2015). "Four principles of bio-musicology." Phil. Trans. R. Soc. B **370**(1664): 20140091.

Fitch, W. T., B. de Boer, et al. (2016). "Monkey vocal tracts are speech-ready." Science Advances **2**(12): e1600723.

Fitch, W. T., B. de Boer, et al. (2017). "Response to Lieberman on "Monkey vocal tracts are speech-ready"." Science Advances **3**(7): e1701859.

Fitch, W. T., M. D. Hauser, et al. (2005). "The evolution of the language faculty: clarifications and implications." Cognition **97**(2): 179-210.

Fitch, W. T., L. Huber, et al. (2010). "Social cognition and the evolution of language: constructing cognitive phylogenies." Neuron **65**(6): 795-814.

Fitch, W. T. and D. Mietchen (2013). "3 Convergence and Deep Homology in the Evolution of Spoken." Birdsong, speech, and language: Exploring the evolution of mind and brain: 45.

Flach, R. (2005). "The transition from synchronization to continuation tapping." Human Movement Science **24**(4): 465-483.

Fodor, J. (1983). The modularity of the mind. Cambrigde, MA, MIT Press.

Fodor, J. (2000). The mind doesn't work that way: The scope and limits of computational psychology. Cambridge, MA, MIT Press.

Fogassi, L., P. F. Ferrari, et al. (2005). "Parietal lobe: from action organization to intention understanding." Science **308**(5722): 662-667.

Fogel, A. (1993). Developing through relationships, University of Chicago Press.

Fogel, A. (1993). "Two principles of communication: Co-regulation and framing." New perspectives in early communicative development: 9-22.

Foley, R. (2004). Music and mosaics: The evolution of human abilities. ESF conference on Music, Language and Human Evolution. Reading, UK.

Foley, R. (2005). The emergence of culture in the context of hominin evolutionary patterns. Evolution and Culture. S. Levinson and P. Jaisson. Cambridge, Massachusetts, MIT Press**:** 53-78.

Foley, R. A. (2012). "Music and Mosaics: The Evolution of Human Abilities." Music, Language, and Human Evolution: 31.

Foley, R. A. (2016). "Mosaic evolution and the pattern of transitions in the hominin lineage." Phil. Trans. R. Soc. B **371**(1698): 20150244.

Foley, R. A. and R. Lewin (2013). Principles of human evolution, John Wiley & Sons.

Folk, C. L., R. W. Remington, et al. (1992). "Involuntary covert orienting is contingent on attentional control settings." Journal of Experimental Psychology Human Perception and Performance **18**: 1030-1030.

Fonagy, I. (1978). "A new method of investigating the perception of prosodic features " Language and Speech **21**: 34-39.

Ford, M. P., R. C. Wagenaar, et al. (2007). "The effects of auditory rhythms and instruction on walking patterns in individuals post stroke." Gait & Posture **26**(1): 150-155.

Formby, C., R. Thomas, et al. (1989). "Regional cerebral blood flow for singers and nonsingers while speaking, singing, and humming a rote passage." Brain and language **36**(4): 690-698.

Foxton, J., J. Dean, et al. (2004). "Characterization of deficits in pitch perception underlying 'tone deafness2." Brain **127**(4): 801-810.

Foxton, J. M., J. L. Dean, et al. (2004). "Characterization of deficits in pitch perception underlying 'tone deafness'." Brain **127**(4): 801-810.

Fraisse, P. (1982). "Rhythm and tempo." The psychology of music **1**: 149-180.

Fraisse, P. (1984). "Perception and estimation of time." Annual review of psychology **35**(1): 1-37.

Franěk, M., J. Mates, et al. (1991). "Finger tapping in musicians and nonmusicians." International Journal of Psychophysiology **11**(3): 277-279.

Franz, E. A., H. N. Zelaznik, et al. (1992). "Evidence of common timing processes in the control of manual, orofacial, and speech movements." Journal of Motor behavior **24**(3): 281-287.

Frayer, D. and C. Nicolay (2000). Fossil evidence for the origin of speech sounds. The origins of music. N. Wallin, B. Merker and S. Brown. Massachusetts, MIT Press**:** 217-234.

Freeman, W. (2000). A neurobiological role of music in social bonding. The origins of music. N. Wallin, B. Merker and S. Brown. Cambridge, MA, MIT Press**:** 411-424.

Freyd, J. J. and J. Johnson (1987). "Probing the time course of representational momentum." Journal of Experimental Psychology: Learning, Memory, and Cognition **13**(2): 259.

Friberg, A. and J. Sundberg (1999). "Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners." The Journal of the Acoustical Society of America **105**(3): 1469-1484.

Friberg, A. and A. Sundströöm (2002). "Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern." Music Perception: An Interdisciplinary Journal **19**(3): 333-349.

Fried, R. and L. Berkowitz (1979). "Music that charms and can influence helpfulness." Journal of Applied Social Psychology **9**: 199-208.

Frijda, N. (1986). The emotions. Studies in emotion and social interaction. New York, Cambridge University Press.

Frijda, N., A. Manstead, et al., Eds. (2000). Emotions and beliefs: How feelings influence thoughts. Cambridge, Cambridge University Press.

Frijda, N. and M. Zeelenberg (2001). Appraisal: What is the dependent? Appraisal processes in emotion. K. Scherer, A. Schorr and T. Johnstone. Oxford, Oxford University Press**:** 141-155.

Fritz, J., M. Mishkin, et al. (2005). "In search of an auditory engram." Proceedings of the National Academy of Sciences of the United States of America **102**(26): 9359-9364.

Fritz, T., S. Jentschke, et al. (2009). "Universal recognition of three basic emotions in music." Current Biology **19**(7): 573-576.

Fu, Y. and W. Li (1996). "Estimating the age of the common ancestor of men from the ZFY intron." Science **272**: 1356-1357; 1360-1361.

Fujii, S., M. Hirashima, et al. (2011). "Synchronization error of drum kit playing with a metronome at different tempi by professional drummers." Music Perception: An Interdisciplinary Journal **28**(5): 491-503.

Fujioka, T., L. Trainor, et al. (2004). "Musical training enhances automatic encoding of melodic contour and intervalic structure." Journal of Cognitive Neuroscience **16**: 1010-1021.

Gabrielsson, A. (1973). "Adjective ratings and dimension analysis of auditory rhythm patterns." Scandanavian Journal of Psychology **14**: 244-260.

Gabrielsson, A. (2001). Emotions in strong experiences with music. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press.

Gabrielsson, A. and E. Lindstroem (2001). The influence of musical structure on emotional response. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press**:** 223-248.

Gagnon, L. and I. Peretz (2003). "Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies." Cognition & Emotion **17**(1): 25-40.

Gale, R. (2016). The philosophy of time: A collection of essays, Springer.

Gallese, V. (2003). "The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity." Psychopathology **36**(4): 171-180.

Gallese, V. (2007). "Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition." Philosophical Transactions of the Royal Society B: Biological Sciences **362**(1480): 659-669.

Gallese, V., C. Keysers, et al. (2004). "A unifying view of the basis of social cognition." Trends in cognitive sciences **8**(9): 396-403.

Gallese, V. and M. Rochat (2013). "The evolution of motor cognition: its role in the development of social cognition and implications for autism spectrum disorder." The infant mind: Origins of the social brain: 19-47.

Gallotti, M., M. Fairhurst, et al. (2017). "Alignment in social interactions." Consciousness and cognition **48**: 253-261.

Gärdenfors, P. (2004). "Cooperation and the Evolution of Symbolic Communication." Evolution of communication systems: 237.

Garrido, M. I., J. M. Kilner, et al. (2009). "The mismatch negativity: a review of underlying mechanisms." Clinical neurophysiology **120**(3): 453-463.

Garrod, S. and M. J. Pickering (2015). "The use of content and timing to predict turn transitions." Frontiers in psychology **6**: 751.

Gehrke, J. (1995). Sensorimotor synchronization: the intensity of afferent feedback affects the timing of movements.

Geissmann, T. (2000). "Duet songs of the siamang, Hylobates syndactylus: I. Structure and organisation." Primate Report **56**: 103-123.

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

Gentilucci, M. and M. C. Corballis (2006). "From manual gesture to speech: A gradual transition." Neuroscience & Biobehavioral Reviews **30**(7): 949-960.

Genty, E., Z. Clay, et al. (2014). "Multi-modal use of a socially directed call in bonobos." PloS one **9**(1): e84738.

Gergely, G., H. Bekkering, et al. (2002). "Developmental psychology: Rational imitation in preverbal infants." Nature **415**(6873): 755.

Gergely, G., Z. Nádasdy, et al. (1995). "Taking the intentional stance at 12 months of age." Cognition **56**(2): 165-193.

Gerloff, C., J. Richard, et al. (1998). "Functional coupling and regional activation of human cortical motor areas during simple, internally paced and externally paced finger movements." Brain: a journal of neurology **121**(8): 1513-1531.

Gerrads-Hesse, A., K. Spies, et al. (1994). "Experimental inductions of emotional states and their effectiveness: a review." British Journal of Psycholog **85**: 55-78.

Gerry, D., A. Unrau, et al. (2012). "Active music classes in infancy enhance musical, communicative and social development." Developmental science **15**(3): 398-407.

Geschwind, N. (1984). "The brain of a learning-disabled individual." Annals of Dyslexia **34**: 319-327.

Gibbon, D. (2003). Computational modelling of rhythm as alternation, iteration and hierarchy. Proceedings of ICPhS.

Gibbon, J. and L. Allan (1984). Timing and time perception, New York, NY: New York Academy of Sciences, 1984.

Gibbon, J., C. Malapani, et al. (1997). "Toward a neurobiology of temporal cognition: advances and challenges." Current opinion in neurobiology **7**(2): 170-184.

Gibson, J. J. (1966). "The senses considered as perceptual systems." Oxford, Houghton Mifflin.

Gilbert, M. (1992). On social facts, Princeton University Press.

Gill, K. Z. and D. Purves (2009). "A biological rationale for musical scales." PloS one **4**(12): e8144.

Gill, S. P. (2007). "Entrainment and musicality in the human system interface." AI & society **21**(4): 567-605.

Gill, S. P. (2012). "Rhythmic synchrony and mediated interaction: towards a framework of rhythm in embodied interaction." AI & society **27**(1): 111-127.

Gill, S. P. (2015). Tacit engagement. Tacit Engagement, Springer**:** 1-34.

Gill, S. P. and J. Borchers (2003). "Knowledge in co-action: social intelligence in collaborative design activity." AI & SOCIETY **17**(3-4): 322-339.

Gilman, B. (1891). "Report of an experimental test of musical expressiveness." American Journal of Psychology **4**: 558-576.

Gisladottir, R. S., D. J. Chwilla, et al. (2015). "Conversation electrified: ERP correlates of speech act recognition in underspecified utterances." PloS one **10**(3): e0120068.

Glennie, E. (1993). "Hearing essay." Published online at www.evelyn.co.uk/ hearing_essay.aspx.

Glory, A. (1964). "La grotte de Roucador." Bulletin de la Societe Prehistorique Francaise **61**: 166-169.

Glory, A. (1965). "Nouvellew decouvertes de dessins rupestres sur le causse de Gramat." Bulletin de la Societe Prehistorique Francaise **62**: 528-536.

Godefroy, O., D. Leys, et al. (1995). "Psychoacoustical deficits related to bilateral subcortical hamorrhages. A case with apperceptive auditory agnosia." Cortex **31**(1): 149-159.

Goebl, W. and C. Palmer (2008). "Tactile feedback and timing accuracy in piano performance." Experimental brain research **186**(3): 471-479.

Goehr, L. (1993). Music has no meaning to speak of: on the politics of musical interpretation. The interpretation of music: Philosophical essays. M. Krausz. Oxford, Oxford University Press.

Gomart, E. and A. Hennion (1999). A sociology of attachment: music amateurs, drug users. Actor network theory and after. J. Law and J. Hazzart. Oxford, Blackwell**:** 220-247.

Good, J. (2008). Linguistic universals and language change, Oxford University Press.

Goodall, J. (1986). "The chimpanzees of Gombe: Patterns of behavior." Belknap Press.

Gopnik, A. (1999). Theory of mind. MIT encyclopedia of cognitive sciences. R. Wilson and F. Keil. Cambrigde, MA, MIT Press**:** 838-841.

Gosselin, N., P. Jolicœur, et al. (2009). "Impaired memory for pitch in congenital amusia." Annals of the New York Academy of Sciences **1169**(1): 270-272.

Gosselin, N., I. Peretz, et al. (2005). "Impaired recognition of scary music following unilateral temporal lobe excision." Brain **128**: 628-640.

Gould, S. and R. Lewontin (1979). "The spandrels of San Marco and the panglossian paradigm." Proceedings of the Royal Society of London B **205**(1161): 581-598.

Gould, S. and E. Vrba (1982). "Exaptation - a missing term in the science of form." Paleobiology **8**: 4-15.

Grabe, E. and E. L. Low (2002). "Durational variability in speech and the rhythm class hypothesis." Papers in laboratory phonology **7**(515-546).

Grahn, J. A. and M. Brett (2007). "Rhythm and beat perception in motor areas of the brain." Journal of cognitive neuroscience **19**(5): 893-906.

Grahn, J. A. and M. Brett (2009). "Impairment of beat-based rhythm discrimination in Parkinson's disease." Cortex **45**(1): 54-61.

Grahn, J. A. and J. D. McAuley (2009). "Neural bases of individual differences in beat perception." Neuroimage **47**(4): 1894-1903.

Grahn, J. A. and J. B. Rowe (2009). "Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception." Journal of Neuroscience **29**(23): 7540-7548.

Grahn, J. A. and J. B. Rowe (2012). "Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity." Cerebral cortex **23**(4): 913-921.

Gratier, M. and E. Devouche (2011). "Imitation and repetition of prosodic contour in vocal interaction at 3 months." Developmental Psychology **47**(1): 67.

Gratier, M., E. Devouche, et al. (2015). "Early development of turn-taking in vocal interaction between mothers and infants." Front. Psychol **6**(1167): 10.3389.

Gratier, M. and C. Trevarthen (2008). "Musical narrative and motives for culture in mother-infant vocal interaction." Journal of Consciousness Studies **15**(10): 122.

Grau-Sánchez, J., J. L. Amengual, et al. (2013). "Plasticity in the sensorimotor cortex induced by Music-supported therapy in stroke patients: a TMS study." Frontiers in human neuroscience **7**.

Greenfield, M. (1994). "Cooperation and conflict in the evolution of signal interactions." Annual Review of Ecological Systems **25**: 97-126.

Grefkes, C., A. Ritzl, et al. (2004). "Human medial intraparietal cortex subserves visuomotor coordinate transformation." Neuroimage **23**(4): 1494-1506.

Gregory, A. (1997). The roles of music in society: The ethnomusicological perspective. The social psychology of music. D. Hargreaves and A. North. Oxford, Oxford University Press**:** 123-140.

Grell, A., J. Sundberg, et al. (2009). "Rapid pitch correction in choir singers." The Journal of the Acoustical Society of America **126**(1): 407-413.

Griffin, Z. M. and K. Bock (2000). "What the eyes say about speaking." Psychological science **11**(4): 274-279.

Griffiths, T., A. Rees, et al. (1997). "Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study." Brain **120**: 785-794.

Griffiths, T., J. Warren, et al. (2004). "“When the feeling’s gone”: a selective loss of musical emotion." Journal of Neurology, Neurosurgery & Psychiatry **75**(2): 344-345.

Gritten, A. and E. King (2006). Music and gesture, Ashgate Publishing, Ltd.

Grondin, S. (2014). About the (non) scalar property for time perception. Neurobiology of interval timing, Springer**:** 17-32.

Gross, J. J. (1998). "The emerging field of emotion regulation: An integrative review." Review of general psychology **2**(3): 271.

Gross, J. J. and R. A. Thompson (2007). "Emotion regulation: Conceptual foundations." NY, Guildford Press.

Grube, M., F. E. Cooper, et al. (2010). "Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration." Proceedings of the National Academy of Sciences **107**(25): 11597-11601.

Grube, M., K.-H. Lee, et al. (2010). "Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals." Frontiers in psychology **1**.

Gruber, D. R. (2017). "Three Forms of Neurorealism: Explaining the Persistence of the “Uncritically Real” in Popular Neuroscience News." Written Communication **34**(2): 189-223.

Gruber, R. P. and R. A. Block (2013). "The flow of time as a perceptual illusion." The Journal of Mind and Behavior: 91-100.

Gruber, T. and D. Grandjean (2016). "A comparative neurological approach to emotional expressions in primate vocalizations." Neuroscience & Biobehavioral Reviews. **73**: 182-190.

Gruber, T. and K. Zuberbühler (2013). "Vocal recruitment for joint travel in wild chimpanzees." PloS one **8**(9): e76073.

Gundlach, R. (1935). "Factors determining the characterization of musical phrases." American Journal of Psychology **47**: 624-644.

Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. Proceedings of Speech Prosody 2002 Conference.

Hafke-Dys, H., A. Preis, et al. (2016). "Violinists' Perceptions of and Motor Reactions to Fundamental Frequency Shifts Introduced in Auditory Feedback." Acta Acustica united with Acustica **102**(1): 155-158.

Hafke-Dys, H. Z., A. Preis, et al. (2013). "Comparison of perceptual and motor responses to changes in intensity and voice fundamental frequency." Acta Acustica united with Acustica **99**(3): 457-464.

Hage, S. R., N. Gavrilov, et al. (2013). "Cognitive control of distinct vocalizations in rhesus monkeys." Journal of Cognitive Neuroscience **25**(10): 1692-1701.

Hage, S. R. and A. Nieder (2013). "Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations." Nature communications **4**: 2409.

Hage, S. R. and A. Nieder (2016). "Dual neural network model for the evolution of speech and language." Trends in neurosciences **39**(12): 813-829.

Hagen, E. and G. Bryant (2003). "Music and dance as a coalition signaling system." Human Nature **14**(1): 21-51.

Hagen, E. and P. Hammerstein (in press). "Did neanderthals and other early humans sing? Seeking the biological roots of music in the territorial advertisements of primates, lions, hyenas, and wolves." Musicae Scientiae, special issue on Music and Evolution. 291-321.

Haggard, P. (2008). "Human volition: towards a neuroscience of will." Nature reviews. Neuroscience **9**(12): 934.

Hahn, J. and S. Muenzel (1995). "Knochenfloeten aus den Aurignacien des Geissenkloesterle bei Blaubeuren, Alb-Donau-Kreis." Fundberichte aus Baden-Wuertemberg **20**: 1-12.

Haidt, J., J. Patrick Seder, et al. (2008). "Hive psychology, happiness, and public policy." The Journal of Legal Studies **37**(S2): S133-S156.

Hain, T. C., T. A. Burnett, et al. (2000). "Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex." Exp Brain Res **130**: 133-141.

Hain, T. C., T. A. Burnett, et al. (2001). "Effects of delayed auditory feedback (DAF) on the pitch-shift reflex." The Journal of the Acoustical Society of America **109**(5): 2146-2152.

Hall, M. (2004). "A review of hypotheses for the functions of avian duetting " Behavioral Ecology and Sociobiology **55**(5): 415-430.

Hall, M. L. (2009). "A review of vocal duetting in birds." Advances in the Study of Behavior **40**: 67-121.

Hallam, S., I. Cross, et al. (2016). Oxford handbook of music psychology, Oxford University Press.

Ham, J. and E. Tronick (2006). "Infant Resilience to the Stress of the Still-Face." Annals of the New York Academy of Sciences **1094**(1): 297-302.

Hamilton, M. J., B. T. Milne, et al. (2007). "The complex structure of hunter–gatherer social networks." Proceedings of the Royal Society of London B: Biological Sciences **274**(1622): 2195-2203.

Hammerschmidt, K., T. Freudenstein, et al. (2001). "Vocal development in squirrel monkeys." Behaviour **138**(9): 1179-1204.

Han, Y., J. Wang, et al. (1999). "Slow tonic muscle fibers in the thyroarytenoid muscles of human vocal folds; a possible specialization for speech." The Anatomical Record **256**(2): 146-157.

Hannon, E. and S. Johnson (2005). "Infants use meter to categorize rhythms and melodies: Implications for musical structure learning." Cognitive Psychology **50**(4): 354-377.

Hannon, E. and S. Trehub (2005a). "Metrical categories in infancy and adulthood." Psychological Science **16**: 48-55.

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

Harcourt-Smith, W. E. (2015). Origin of bipedal locomotion. Handbook of Paleoanthropology, Springer**:** 1919-1959.

Hare, B., E. Addessi, et al. (2003). "Do capuchin monkeys, Cebus apella, know what conspecifics do and do not see?" Animal Behaviour **65**(1): 131-142.

Hare, B., J. Call, et al. (2000). "Chimpanzees know what conspecifics do and do not see." Animal Behaviour **59**(4): 771-785.

Hare, B., J. Call, et al. (2001). "Do chimpanzees know what conspecifics know?" Animal Behaviour **61**(1): 139-151.

Hargreaves, D. and A. North (1999). "The functions of music in everyday life: Redefining the social in music psychology " Psychology of Music **27**(1): 71-83.

Harrington, D. L., L. A. Boyd, et al. (2004). "Neural representation of interval encoding and decision making." Cognitive Brain Research **21**(2): 193-205.

Harrington, F. and L. Mech (1983). "Wolf pack spacing: Howling as a territory-independent spacing mechanism in a territorial population." Behavioral Ecology and Sociobiology **12**(2): 161-168.

Harter, S. (1981). "A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components." Developmental psychology **17**(3): 300.

Hary, D. and G. Moore (1987). "Synchronizing human movement with an external clock source." Biological cybernetics **56**(5): 305-311.

Hary, D. and G. P. Moore (1985). "Temporal tracking and synchronization strategies." Human Neurobiology **4**(2): 73-79.

Hasegawa, A., K. Okanoya, et al. (2011). "Rhythmic synchronization tapping to an audio–visual metronome in budgerigars." Scientific reports **1**: 120.

Hattori, Y., M. Tomonaga, et al. (2013). "Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee." Scientific reports **3**: 1566.

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

Hauser, M. and J. McDermott (2003). "The evolution of the music faculty: A comparative perspective." Nature Neuroscience **6**(7): 663-668.

Hauser, M. and J. McDermott (2006). "thought on an empirical approach to the the evolutionary origins of music." Music Perception **24**(1): 111-115.

Hayden, R., A. A. Clair, et al. (2009). "The effect of rhythmic auditory stimulation (RAS) on physical therapy outcomes for patients in gait training following stroke: a feasibility study." International Journal of Neuroscience **119**(12): 2183-2195.

He, C. and L. J. Trainor (2009). "Finding the pitch of the missing fundamental in infants." Journal of Neuroscience **29**(24): 7718-8822.

Heaton, P. (2009). "Assessing musical skills in autistic children who are not savants." Philosophical Transactions of the Royal Society of London B: Biological Sciences **364**(1522): 1443-1447.

Heaton, P., R. E. Davis, et al. (2008). "Research note: Exceptional absolute pitch perception for spoken words in an able adult with autism." Neuropsychologia **46**(7): 2095-2098.

Hebert, S., A. Racette, et al. (2003). "Revisiting the dissociation between singing and speaking in expressive aphasia." Brain **126**: 1838-1850.

Heinsohn, R. (1997). "Group territoriality in two populations of African lions." Animal Behaviour **53**(6): 1143-1147.

Heldner, M. and J. Edlund (2010). "Pauses, gaps and overlaps in conversations." Journal of phonetics **38**(4): 555-568.

Hellmer, K. and G. Madison (2015). "Quantifying microtiming patterning and variability in drum kit recordings." Music Perception: An Interdisciplinary Journal **33**(2): 147-162.

Henry, L., A. Craig, et al. (2015). "Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model." Frontiers in psychology **6**: 1416.

Henry, M. J. and B. Herrmann (2014). "Low-frequency neural oscillations support dynamic attending in temporal context." Timing & Time Perception **2**(1): 62-86.

Henry, M. J. and J. D. McAuley (2010). "On the prevalence of congenital amusia." Music Perception: An Interdisciplinary Journal **27**(5): 413-418.

Henry, M. J. and J. D. McAuley (2013). "Failure to apply signal detection theory to the Montreal Battery of Evaluation of Amusia may misdiagnose amusia." Music Perception: An Interdisciplinary Journal **30**(5): 480-496.

Herbinger, I., C. Boesch, et al. (2001). "Territory characteristics among three neighbouring chimpanzee communities in Tai National Park, Cote d'Ivoire." International Journal of Primatology **22**(2): 143-167.

Hesslow, G. (2012). "The current status of the simulation theory of cognition." Brain research **1428**: 71-79.

Hewitt, G., A. Maclarnon, et al. (2002). "The functions of layrngeal air sacs in primates: a new hypothesis." Folia Primatologia **73**: 70-94.

Hihara, S., H. Yamada, et al. (2003). "Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys." Neuroscience Research **45**(4): 383-389.

Himberg, T. (2014). Interaction in musical time, University of Cambridge.

Himberg, T., L. Hirvenkari, et al. (2015). "Word-by-word entrainment of speech rhythm during joint story building." Frontiers in psychology **6**.

Hinde, R. (1975). "The comparative study of non-verbal communication." The body as a medium of expression: 107-142.

Hinde, R. A. (1974). Biological bases of human social behaviour, McGraw-Hill.

Hirel, C., Y. Lévêque, et al. (2014). "Acquired amusia and musical anhedonia." Revue neurologique **170**(8-9): 536-540.

Hirsh, I. J. and C. E. Sherrick Jr (1961). "Perceived order in different sense modalities." Journal of experimental psychology **62**(5): 423.

Hockett, C. F. (1959). "Animal" languages" and human language." Human Biology **31**(1): 32-39.

Hockett, C. F. and S. A. Altmann (1968). "A note on design features." Animal communication: Techniques of study and results of research: 61-72.

Hockett, C. F. and C. D. Hockett (1960). "The origin of speech." Scientific American **203**(3): 88-97.

Hoequist Jr, C. (1983). "Durational correlates of linguistic rhythm categories." Phonetica **40**(1): 19-31.

Hoeschele, M., R. G. Weisman, et al. (2013). "Chickadees fail standardized operant tests for octave equivalence." Animal cognition **16**(4): 599-609.

Hoeschele, M., R. G. Weisman, et al. (2012). "Pitch chroma discrimination, generalization, and transfer tests of octave equivalence in humans." Attention, Perception, & Psychophysics **74**(8): 1742-1760.

Hofer, H. and M. East (1993). "The commuting system of the Serengeti spotted hyaenas: How a predator copes with migratory prey. I." Animal Behaviour **46**(1): 547-557.

Hofer, M. (1990). Early symbiotic processes: Hard evidence from a soft place. Pleasure beyond the pleasure principle. R. Glick and S. Bones. New Haven, Yale University Press**:** 55-78.

Hofer, M. A. (1995). "Hidden regulators." Attachment theory: Social, developmental and clinical perspectives: 203-230.

Hoffding, S. (2014). "What is Skilled Coping?: Experts on Expertise." Journal of Consciousness Studies **21**(9-10): 49-73.

Hofstad, E., Y. Espmark, et al. (2002). "The relationship between song performance and male quality in snow buntings (Plectrophenax nivalis) " Canadian Journal of Zoology **80**(3): 524-531.

Hohmann, G. and B. Fruth (1994). "Structure and use of distance calls in wild bonobos (Pan paniscus)." International Journal of Primatology **15**(5): 767-782.

Holler, J., Kendrick, K. H., Casillas, M., and S. C. Levinson (2015). "Editorial: Turn-Taking in Human Communicative Interaction." Frontiers in psychology **6**: doi:10.3389/3978-3382-88919-88825-88912.

Holloway, R. L. (1983). "Human paleontological evidence relevant to language behavior." Human neurobiology **2**(3): 105-114.

Honeycutt, J. and M. Eidenmueller (2001). Communication and attribution: an exploration of the effects of music and mood on intimate couples' verbal and noverbal conflict resolution behaviours. Attribution, communication behaviour, and close relationships: Advances in personal relations. V. Manusov and J. Harvey. New York, Cambridge University Press**:** 21-37.

Honing, H. (2014). Musical cognition: a science of listening, Transaction Publishers.

Honing, H., H. Merchant, et al. (2012). "Rhesus monkeys (Macaca mulatta) detect rhythmic groups in music, but not the beat." PloS one **7**(12): e51369.

Honing, H. and A. Ploeger (2012). "Cognition and the evolution of music: Pitfalls and prospects." Topics in Cognitive Science **4**(4): 513-524.

Hopkins, W. D., J. P. Taglialatela, et al. (2007). "Chimpanzees differentially produce novel vocalizations to capture the attention of a human." Animal Behaviour **73**(2): 281-286.

Hosokawa, T., R. Nakamura, et al. (1981). "Monotic and dichotic fusion thresholds in patients with unilateral subcortical lesions." Neuropsychologia **19**(2): 241-248.

Houtsma, A. J. and J. L. Goldstein (1972). "The central origin of the pitch of complex tones: Evidence from musical interval recognition." The Journal of the Acoustical Society of America **51**(2B): 520-529.

Houtsma, A. J. and J. Smurzynski (1990). "Pitch identification and discrimination for complex tones with many harmonics." The Journal of the Acoustical Society of America **87**(1): 304-310.

Hove, M. J. (2008). "Shared circuits, shared time, and interpersonal synchrony." Behavioral and Brain Sciences **31**(1): 29-30.

Hove, M. J. and J. L. Risen (2009). "It's all in the timing: Interpersonal synchrony increases affiliation." Social cognition **27**(6): 949-960.

Hulse, S. H. and C. L. Kline (1993). "The perception of time relations in auditory tempo discrimination." Animal Learning & Behavior **21**(3): 281-288.

Hulse, S. H. and S. C. Page (1988). "Toward a comparative psychology of music perception." Music Perception: An Interdisciplinary Journal **5**(4): 427-452.

Hulse, S. H., A. H. Takeuchi, et al. (1992). "Perceptual invariances in the comparative psychology of music." Music Perception: An Interdisciplinary Journal **10**(2): 151-184.

Huppert, F. A., N. Baylis, et al. (2005). The science of well-being, Oxford University Press, USA.

Hurford, J. R. (2003). "The language mosaic and its evolution." Studies in the Evolution of Language **3**: 38-57.

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

Huron, D. (2009). "Why some ethnomusicologists don’t like music cognition: Finding common ground in the study of musical minds." Proceedings of the European Society for the Cognitive Sciences of Music. Finland: Jyväskylä.

Huron, D., D. Kinney, et al. (2006). "Influence of pitch height on the perception of submissiveness and threat in musical passages." Empirical Musicology Review **1**(3): 170-177.

Huron, D. B. (2006). Sweet anticipation: Music and the psychology of expectation, MIT press.

Hurt, C. P., R. R. Rice, et al. (1998). "Rhythmic auditory stimulation in gait training for patients with traumatic brain injury." Journal of Music Therapy **35**(4): 228-241.

Husain, G., W. Thompson, et al. (2002). "Effects of musical tempo and mode on arousal, mood, and spatial abilities." Music Perception **20**: 151-171.

Hutchins, S. and S. Moreno (2013). "The Linked Dual Representation model of vocal perception and production." Frontiers in psychology **4**.

Hutchins, S. and I. Peretz (2013). "Vocal pitch shift in congenital amusia (pitch deafness)." Brain and language **125**(1): 106-117.

Hutchins, S., C. Roquet, et al. (2012). "The vocal generosity effect: How bad can your singing be?" Music Perception: An Interdisciplinary Journal **30**(2): 147-159.

Hutchins, S., J. M. Zarate, et al. (2010). "An acoustical study of vocal pitch matching in congenital amusia." The Journal of the Acoustical Society of America **127**(1): 504-512.

Hutchins, S. M. and I. Peretz (2012). "A frog in your throat or in your ear? Searching for the causes of poor singing." Journal of Experimental Psychology: General **141**(1): 76.

Huxley, J. (1914). "The courtship habits of the Great Crested Grebe (Podiceps cristatus) together with a discussion of the evolution of courtship in birds." Journal of the Linnean Society of London, Zoology **53**: 253-292.

Huyge, D. (1990). "Mousterian skiffle? Note on a middle Paleolithic engraved bone from Schulen, Belgium." Rock Art Research **7**: 125-132.

Hyde, K. and I. Peretz (2004). "Brains that are out-of-tune but in-time." Psychological Science**15**: 356-60.

Hyde, K. L., R. J. Zatorre, et al. (2010). "Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia." Cerebral Cortex **21**(2): 292-299.

Ilie, G. and W. Thompson (2006). "A comparison of acoustic cues in music and speech for three dimensions of affect." Music Perception **23**(4): 319-329.

Indefrey, P. (2011). "The spatial and temporal signatures of word production components: a critical update." Frontiers in psychology **2**.

Indefrey, P. and W. J. Levelt (2004). "The spatial and temporal signatures of word production components." Cognition **92**(1): 101-144.

Itoh, K., S. Suwazono, et al. (2010). "Central auditory processing of noncontextual consonance in music: an evoked potential study." The Journal of the Acoustical Society of America **128**(6): 3781-3787.

Iversen, S. D. and P. J. Fray (1982). Brain catecholamines in relation to affect. The neural basis of behavior, Springer**:** 229-269.

Ivry, R. B. and R. E. Hazeltine (1995). "Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism." Journal of experimental psychology: Human perception and performance **21**(1): 3.

Ivry, R. B. and T. C. Richardson (2002). "Temporal control and coordination: the multiple timer model." Brain and cognition **48**(1): 117-132.

Ivry, R. B. and J. E. Schlerf (2008). "Dedicated and intrinsic models of time perception." Trends in cognitive sciences **12**(7): 273-280.

Ivry, R. B. and R. M. Spencer (2004). "The neural representation of time." Current opinion in neurobiology **14**(2): 225-232.

Iyer, V. (2002). "Embodied mind, situated cognition, and expressive microtiming in African-American music." Music Perception: An Interdisciplinary Journal **19**(3): 387-414.

Izumi, A. (2000). "Japanese monkeys perceive sensory consonance of chords." The Journal of the Acoustical Society of America **108**(6): 3073-3078.

Jablonski, N. and L. Aiello, Eds. (1998). The origin and diversification of language. Wattis Symposium Series in A. San Francisco, California Academy of Sciences.

Jackendoff, R. and S. Pinker (2005). "The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky)." Cognition **97**(2): 211-225.

Jackendorff, R. (1994). Patterns in the mind: Language and human nature. New York, Basic Books.

Jackson, J. C., J. Jong, et al. (2018). "Synchrony and Physiological Arousal Increase Cohesion and Cooperation in Large Naturalistic Groups." Scientific reports **8**(1): 127.

Jackson, L., R. Heffner, et al. (1999). "Free-field audiogram of the Japanese macaque (Macaca fuscata)." Journal of the Acoustical Society of America **106**(5): 3017-3023.

Jaffe, J. and S. Anderson (1979). Communication rhythms and the evolution of language. Of speech and temporal speech patterns in interpersonal contexts. A. Siegman and S. Feldstein. Hillsdale NJ, Erlbaum Associates.

Jaffe, J., B. Beebe, et al. (2001). "Rhythms of dialogue in infancy: Coordinated timing in development." Monographs of the society for research in child development: i-149.

James, W. (1890). 1950. The Principles of Psychology, Vol. 1, Dover: New York.

Jamieson, I. (1986). "The functional approach to behavior: Is it useful." The American Naturalist **127**: 195-208.

Janata, P. (2001). Neurophysiological mechanisms underlying auditory image formation in music. In RI God¯ y & H. J¯rgensen (Eds.), Musical Imagery (pp. 27-47). Lisse: Swets & Zeitlinger, Citeseer.

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

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

Jäncke, L., R. Loose, et al. (2000). "Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli." Cognitive Brain Research **10**(1): 51-66.

Janik, V. M. and P. J. Slater (1997). "Vocal learning in mammals." Advances in the Study of Behaviour **26**: 59-100.

Jantzen, K., F. Steinberg, et al. (2005). "Functional MRI reveals the existence of modality and coordination-dependent timing networks." Neuroimage **25**(4): 1031-1042.

Jaquet, L., B. Danuser, et al. (2014). "Music and felt emotions: How systematic pitch level variations affect the experience of pleasantness and arousal." Psychology of Music **42**(1): 51-70.

Järvinen-Pasley, A., J. Pasley, et al. (2008). "Is the linguistic content of speech less salient than its perceptual features in autism?" Journal of autism and developmental disorders **38**(2): 239-248.

Järvinen-Pasley, A., G. L. Wallace, et al. (2008). "Enhanced perceptual processing of speech in autism." Developmental science **11**(1): 109-121.

Jarvis, E. D. (2004). "Learned birdsong and the neurobiology of human language." Annals of the New York Academy of sciences **1016**(1): 749-777.

Jarvis, E. D. (2007). "Neural systems for vocal learning in birds and humans: a synopsis." Journal of Ornithology **148**(1): 35-44.

Jarvis, E. D. (2010). "Bird song systems: Evolution." In Squire L. (Ed.) Encyclopedia of Neuroscience **2**: 217-19.

Jarvis, E. D. (2013). "Evolution of brain pathways for vocal learning in birds and humans." Birdsong, speech, and language: Exploring the evolution of mind and brain: 63-107.

Jarvis, E. D. (2017). "Birdsong and the Neurobiology of Language." Annual Review of Linguistics **3**(1).

Jeffries, K., J. B. Fritz, et al. (2003). "Words in melody: an H215O PET study of brain activation during singing and speaking." Neuroreport **14**(5): 749-754.

Jentschke, S., S. Koelsch, et al. (2008). "Children with specific language impairment also show impairment of music-syntactic processing." Journal of Cognitive Neuroscience **20**(11): 1940-1951.

Jiang, C., V. K. Lim, et al. (2013). "Difficulties with pitch discrimination influences pitch memory performance: evidence from congenital amusia." PloS one **8**(10): e79216.

Joffe, M. (1997). "Social pressures have selected for an extended juvenile period in primates." Journal of Human Evolution **32**(6): 593-605.

Johanson, D. C. and B. Edgar (1996). From Lucy to language, Simon and Schuster.

Johansson, S. (2005). Origins of language: Constraints on hypothesis. Amsterdam/Philidelphia, John Benjamins Publishing Company.

Johnson, M. H., R. Griffin, et al. (2005). "The emergence of the social brain network: Evidence from typical and atypical development." Development and psychopathology **17**(3): 599-619.

Johnstone, R. (1997). The evolution of animal signals. Behavioural Ecology: An evolutionary approach. J. Krebs and N. Davies. Oxford, Blackwell.

Jones, J. A. and D. Keough (2008). "Auditory-motor mapping for pitch control in singers and nonsingers." Experimental Brain Research **190**(3): 279-287.

Jones, J. A. and K. Munhall (2005). "Remapping auditory-motor representations in voice production." Current Biology **15**(19): 1768-1772.

Jones, J. A. and K. G. Munhall (2000). "Perceptual calibration of F 0 production: Evidence from feedback perturbation." The Journal of the Acoustical Society of America **108**(3): 1246-1251.

Jones, J. A., N. Scheerer, et al. (2013). The relationship between vocal pitch feedback error and event-related brain potentials. Proceedings of Meetings on Acoustics ICA2013, ASA.

Jones, J. L., C. Zalewski, et al. (2009). "Widespread auditory deficits in tune deafness." Ear and hearing **30**(1): 63.

Jones, M. (1976). "Time, our lost dimension: Toward a new theory of perception, attention, and memory." Psychological Review **83**: 325-335.

Jones, M. R. (2009/2016). "Musical time." In Hallam S., Cross., I. & Thaut, M. (Eds.) The handbook of music psychology: 81-92. Oxford University Press.

Jones, M. R. and M. Boltz (1989). "Dynamic attending and responses to time." Psychological review **96**(3): 459.

Jones, M. R. and J. D. Mcauley (2005). "Time judgments in global temporal contexts." Attention, Perception, & Psychophysics **67**(3): 398-417.

Jonides, J. and S. Yantis (1988). "Uniqueness of abrupt visual onset in capturing attention." Perception & Psychophysics **43**(4): 346-354.

Jordania, J. (2006). "Who asked the first question." The origins of human choral singing, intelligence, language and speech. Tblisi, Georgia: Logos.

Jungblut, M., W. Huber, et al. (2012). "The impact of rhythm complexity on brain activation during simple singing: an event-related fMRI study." Restorative neurology and neuroscience **30**(1): 39-53.

Jürgens, U. (1976). "Projections from the cortical larynx area in the squirrel monkey." Experimental Brain Research **25**(4): 401-411.

Jürgens, U. (1994). "The role of the periaqueductal grey in vocal behaviour." Behavioural brain research **62**(2): 107-117.

Jürgens, U. (2002). "Neural pathways underlying vocal control." Neuroscience & Biobehavioral Reviews **26**(2): 235-258.

Jürgens, U. (2009). "The neural control of vocalization in mammals: a review." Journal of Voice **23**(1): 1-10.

Jusczyk, P. and C. Krumhansl (1993). "Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure." Journal of Experimental Psychology: Human Perception and Performance **19**: 627-640.

Juslin, P. (1995). Emotional communication in music viewed through a Brunswikian lens. Music and expression: Proceedings of the Conference of DGM and ESCOM, Bremen, 1995. G. Kleinen. Bremen, DE, University of Bremen**:** 21-25.

Juslin, P. (2000). "Cue utilization in communication of emotion in music performance: Relating performance to perception." Journal of Experimental Psychology: Human Perception and Performance **26**: 1797-1813.

Juslin, P. (2001). Communicating emotion in music performance: a review and a theoretical framework. Music and emotion: theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press.

Juslin, P. (2005). From mimesis to catharsis: expression, perception, and induction of emotion in music. Musical Communication. D. Miell, R. MacDonald and D. Hargreaves. Oxford, Oxford University Press.

Juslin, P. and P. Laukka (2003). "Communication of emotions in vocal expression and music performance: Different channels, same code?" Psychological Bulletin **129**: 770-814.

Juslin, P. and P. Laukka (2004). "Expression, perception, and induction of musical emotion: a review and a questionnaire study of everyday listening." Journal of New Music Research **33**(3): 216-237.

Juslin, P. N. (2016). "Emotional Reactions to Music." In Hallam S., Cross., I. & Thaut, M. (Eds.) The handbook of music psychology: 197. Oxford University Press.

Juslin, P. N. and E. Lindström (2010). "Musical expression of emotions: Modelling listeners' judgements of composed and performed features." Music Analysis **29**(1-3): 334-364.

Juslin, P. N. and J. Sloboda (2011). Handbook of music and emotion: Theory, research, applications, Oxford University Press.

Juslin, P. N. and J. A. Sloboda (2013). "Music and emotion." The psychology of music **3**: 583-645.

Juslin, P. N. and D. Västfjäll (2008). "Emotional responses to music: The need to consider underlying mechanisms." Behavioral and Brain Sciences **31**(5): 559-575.

Justus, T. and Hutsler (2005). "Fundamental issues in the evolutionary psychology of music: Assessing Innateness and Domain Speificity." Music Perception **23**(1): 1-27.

Kaas, J. H., T. A. Hackett, et al. (1999). "Auditory processing in primate cerebral cortex." Current opinion in neurobiology **9**(2): 164-170.

Kalakoski, V. (2001). "Musical imagery and working memory." In Godoy, R. & Jorgenson, H. (Eds.) Musical imagery: 43-56. Taylor and Francis.

Kallinen, K. (2004). Emotion-related psychophysiological response to listening to music wit eyes-open versus eyes-closed: Electrodermal, electrocardiac, and electromyographic measures. Proceedings of the 8th international conference on music perception and cognition. S. Lipscomb, R. Ashley, P. Gjerdingen and P. Webster. Evanston, IL, ICMPC**:** 299-301.

Kaminski, J., J. Call, et al. (2008). "Chimpanzees know what others know, but not what they believe." Cognition **109**(2): 224-234.

Karmiloff-Smith, A. (1995). Beyond modularity: A developmental perspective on cognitive science, MIT press.

Kay, R., M. Cartmill, et al. (1998). "The hypoglossal canal and the origin of human vocal behaviour." Proceedings of the National Academy of Sciences **95**: 5417-5419.

Keele, S. W., R. A. Pokorny, et al. (1985). "Do perception and motor production share common timing mechanisms: A correlational analysis." Acta psychologica **60**(2): 173-191.

Keil, C. (1979). "Tiv Song – the sociology of art in a classless society". University of Chicago Press.

Keller, P. E. (2008). "Joint action in music performance." Emerging Communication **10**: 205.

Keller, P. E. (2008). Joint action in music performance. Emerging communication: Studies on new technologies and practices in communication: Vol. 10. Enacting intersubjectivity: A cognitive and social perspective on the study of interactions. A. C. F. Morganti, & G. Riva. Amsterdam, IOS Press**:** 205-221.

Keller, P. E., G. Novembre, et al. (2014). "Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination." Phil. Trans. R. Soc. B **369**(1658): 20130394.

Keller, P. E., G. Novembre, et al. (2016). "14 Musical Ensemble Performance: Representing Self, Other and Joint Action Outcomes." Shared representations: Sensorimotor foundations of social life: 280.

Kelso, J. S. and B. Kay (1987). "Information and control: A macroscopic analysis of perception-action coupling." Perspectives on perception and action **1**: 3-32.

Keltner, D. and J. Gross (1999). "Functional accounts of emotions." Cognition and Emotion **13**: 465-466.

Kendon, A. (2004). Gesture: Visible action as utterance, Cambridge University Press.

Keough, D., C. Hawco, et al. (2013). "Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback." BMC neuroscience **14**(1): 25.

Kessler, E. J., C. Hansen, et al. (1984). "Tonal schemata in the perception of music in Bali and in the West." Music Perception: An Interdisciplinary Journal **2**(2): 131-165.

Ketal, R. (1975). "Affect, mood, emotion, and feeling: semantic considerations." American Journal of Psychiatry **132**: 1215-1217.

Keverne, E. B., N. D. Martensz, et al. (1989). "Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships." Psychoneuroendocrinology **14**(1-2): 155-161.

Kirschner, S. and B. Ilari (2014). "Joint drumming in Brazilian and German preschool children: Cultural differences in rhythmic entrainment, but no prosocial effects." Journal of Cross-Cultural Psychology **45**(1): 137-166.

Kirschner, S. and M. Tomasello (2009). "Joint drumming: social context facilitates synchronization in preschool children." Journal of experimental child psychology **102**(3): 299-314.

Kirschner, S. and M. Tomasello (2010). "Joint music making promotes prosocial behavior in 4-year-old children." Evolution and Human Behavior **31**(5): 354-364.

Kivy, P. (1991). Music alone: Philosophical reflections on the purely musical experience, Cornell University Press.

Kleber, B., N. Birbaumer, et al. (2007). "Overt and imagined singing of an Italian aria." Neuroimage **36**(3): 889-900.

Klein, M. and W. Hartmann (1981). "Binaural edge pitch." The Journal of the Acoustical Society of America **70**(1): 51-61.

Kleinen, G. (1968). Experimentelle Studien zum musikalischen Ausdruck. Hamburg, DE, Universitaet Hamburg.

Kleinginna, P. and A. Kleinginna (1981). "A categorized list of emotion definitions, with a suggestion for a consensual definition." Motivation and Emotion **5**: 345-371.

Knight, A., P. A. Underhill, et al. (2003). "African Y chromosome and mtDNA divergence provides insight into the history of click languages." Current Biology **13**(6): 464-473.

Knight, C. (1999). "Sex and language as pretend-play." The evolution of culture: An interdisciplinary view: 228-247.

Knight, R.-A. (2011). "Assessing the temporal reliability of rhythm metrics." Journal of the International Phonetic Association **41**(3): 271-281.

Kobinata, N., Ueno, M., Imanishi, Y., & Yoshikawa, H. (2016). Immediate effects of rhythmic auditory stimulation on gait in stroke patients in relation to the lesion site. Journal of physical therapy science **28**(9), 2441-2444.

Kochetkova, V. (1978). Paleoneurology. Washington, Winston & Sons.

Koda, H., M. Basile, et al. (2013). "Validation of an auditory sensory reinforcement paradigm: Campbell's monkeys (Cercopithecus campbelli) do not prefer consonant over dissonant sounds." Journal of Comparative Psychology **127**(3): 265.

Koda, H., C. Oyakawa, et al. (2007). "Experimental evidence for the volitional control of vocal production in an immature gibbon." Behaviour **144**(6): 681-692.

Koehne, S., M. J. Schmidt, et al. (2016). "The role of interpersonal movement synchronisation in empathic functions: Insights from Tango Argentino and Capoeira." International Journal of Psychology **51**(4): 318-322.

Koelsch, S. (2005). "Investigating emotion with music: Neuroscientific approaches." Annals of the New York Academy of Sciences **1060**: 412-418.

Koelsch, S. (2011). "Toward a neural basis of music perception–a review and updated model." Frontiers in psychology **2**.

Koelsch, S. (2012). Brain and music, John Wiley & Sons.

Koelsch, S. (2014). "Brain correlates of music-evoked emotions." Nature Reviews Neuroscience **15**(3): 170-180.

Koelsch, S., T. Fritz, et al. (2005). "Adults and children processing music: an fMRI study." Neuroimage **25**: 1068-1076.

Koelsch, S., T. Fritz, et al. (2006). "Investigating emotion with music: an fMRI study." Human Brain Mapping **27**(3): 239-250.

Koelsch, S., T. Grossmann, et al. (2006). "Children processing music: electric brain responses reveal musical competence and gender differences." Children **15**(5).

Koelsch, S., M. Rohrmeier, et al. (2013). "Processing of hierarchical syntactic structure in music." Proceedings of the National Academy of Sciences **110**(38): 15443-15448.

Kojima, S. (1990). "Comparison of auditory functions in the chimpanzee and human." Folia Primatologica **55**(2): 62-72.

Kojima, S. (1990). "Comparison of auditory functions in the chimpanzee and human." Folia Primatologia **55**(2): 62-72.

Kolers, P. A. and J. M. Brewster (1985). "Rhythms and responses." Journal of experimental psychology: Human perception and performance **11**(2): 150-167.

Komeilipoor, N., M. W. Rodger, et al. (2015). "(Dis-) Harmony in movement: effects of musical dissonance on movement timing and form." Experimental brain research **233**(5): 1585-1595.

Konner, M. (1974). Aspects of the developmental ethology of a foraging people. Ethological studies of child behaviour. N. G. B. Jones, Cambridge University Press**:** 285-304.

Konner, M. (2014). "Conflict and Evolution." Ancestral Landscapes in Human Evolution: Culture, Childrearing and Social Wellbeing: 257.

Krachun, C., J. Call, et al. (2010). "A new change-of-contents false belief test: Children and chimpanzees compared." International Journal of Comparative Psychology **23**(2).

Krachun, C., M. Carpenter, et al. (2009). "A competitive nonverbal false belief task for children and apes." Developmental science **12**(4): 521-535.

Krantz, G., G. Kreutz, et al. (2011). "Bodily Movements Influence HeartRate Variability (HRV) Responses toIsolated Melodic Intervals." Music and medicine **3**(2): 108-113.

Krause, V., B. Pollok, et al. (2010). "Perception in action: the impact of sensory information on sensorimotor synchronization in musicians and non-musicians." Acta psychologica **133**(1): 28-37.

Kronman, U. and J. Sundberg (1987). "Is the musical ritard an allusion to physical motion." Action and perception in rhythm and music **55**: 57-68.

Krueger, J. (2013). "Empathy, enaction, and shared musical experience: evidence from infant cognition." The Emotional Power of Music: Multidisciplinary Perspectives on Musical Arousal, Expression, and Social Control: 177.

Krumhansl, C. (1990). Cognitive foundations of musical pitch. New York, Oxford University Press.

Krumhansl, C. (1997). "An exploratory study of musical emotions and psychophysiology." Canadian Journal of Experimental Psychology **51**: 336-352.

Krumhansl, C. and P. Jusczyk (1990). "Infants' perception of phrase structure in music." Psychological Science **1**: 70-73.

Krumhansl, C. and F. Keil (1982). "Acquisition of the hierarchy of tonal functions in music." Memory & Cognition **10**: 243-251.

Krumhansl, C., P. Toivanen, et al. (2000). "Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks." Cognition **76**(1): 13-58.

Krumhansl, C. L. (1985). "Perceiving Tonal Structure in Music: The complex mental activity by which listeners distinguish subtle relations among tones, chords, and keys in Western tonal music offers new territory for cognitive psychology." American Scientist **73**(4): 371-378.

Krumhansl, C. L. (2000). "Rhythm and pitch in music cognition." Psychological bulletin **126**(1): 159.

Krumhansl, C. L. and L. L. Cuddy (2010). A theory of tonal hierarchies in music. Music perception, Springer New York**:** 51-87.

Krupenye, C., F. Kano, et al. (2016). "Great apes anticipate that other individuals will act according to false beliefs." Science **354**(6308): 110-114.

Kubovy, M. and F. P. Howard (1976). "Persistence of a pitch-segregating echoic memory." Journal of experimental psychology: Human perception and performance **2**(4): 531.

Kühl, O. (2008). Musical semantics, Peter Lang.

Kuhl, P. (1988). "Auditory perception and the evolution of speech." Human Evolution **3**: 19-43.

Kuhl, P. K., R. R. Ramírez, et al. (2014). "Infants' brain responses to speech suggest analysis by synthesis." Proceedings of the National Academy of Sciences **111**(31): 11238-11245.

Kuhlmeier, V., K. Wynn, et al. (2003). "Attribution of dispositional states by 12-month-olds." Psychological science **14**(5): 402-408.

Kung, S.-J., J. L. Chen, et al. (2013). "Interacting cortical and basal ganglia networks underlying finding and tapping to the musical beat." Journal of cognitive neuroscience **25**(3): 401-420.

Ladd, D. R. (2008). Intonational phonology, Cambridge University Press.

Ladefoged, P. l975. A course in phonetics, New York: Harcourt Brace Jovanovich.

Ladwig, J. C., T. C. Prado, et al. (2016). "Facilitating movement performance in cerebral palsy: The impact of rhythmic auditory cueing in a goal-directed reaching task." Journal of Exercise, Movement, and Sport **48**(1): 27.

Lahav, A., E. Saltzman, et al. (2007). "Action representation of sound: audiomotor recognition network while listening to newly acquired actions." Journal of Neuroscience **27**(2): 308-314.

Lahr, M. and R. Foley (1994). "Multiple dispersals and modern human origins." Evolutionary Anthropolgy **3**(2): 48-60.

Laignel-Lavastine, M. and T. Alajouanine (1921). "Un cas d'agnosie auditive." Societe de Neurologie **37**: 194-198.

Laitman, J. (1984). "The anatomy of human speech." Natural History **93**: 20-27.

Laitman, J. and R. Heimbuch (1982). "The basicranium of Plio-Pleistocene hominids as an indicator of their upper respiratory systems." American Journal of Physical Anthropology **59**: 323-344.

Laitman, J., R. Heimbuch, et al. (1979). "The basicranium of fossil hominids as an indicator of their upper respiratory systems." American Journal of Physical Anthropology **51**: 13-34.

Laitman, J. and J. Reidenberg (1988). "Advances in understanding the relationship between the skull base and larynx with comments on the origins of speech." Human Evolution **3**: 99-109.

Lakatos, P., G. Karmos, et al. (2008). "Entrainment of neuronal oscillations as a mechanism of attentional selection." Science **320**(5872): 110-113.

Lakin, J. L. and T. L. Chartrand (2003). "Using nonconscious behavioral mimicry to create affiliation and rapport." Psychological science **14**(4): 334-339.

Lameira, A. R. (2014). "The forgotten role of consonant-like calls in theories of speech evolution." Behavioral and Brain Sciences **37**(6): 559-560.

Lameira, A. R., M. E. Hardus, et al. (2015). "Speech-like rhythm in a voiced and voiceless orangutan call." PloS one **10**(1): e116136.

Lameira, A. R., M. E. Hardus, et al. (2013). "Orangutan (Pongo spp.) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension." The Journal of the Acoustical Society of America **134**(3): 2326-2335.

Lameira, A. R., M. E. Hardus, et al. (2016). "Vocal fold control beyond the species-specific repertoire in an orang-utan." Scientific reports **6**.

Lameira, A. R., I. Maddieson, et al. (2014). "Primate feedstock for the evolution of consonants." Trends in cognitive sciences **18**(2): 60-62.

Lamont, A. and I. Cross (1994). "Children's cognitive representations of musical pitch." Music Perception **12**: 27-55.

Lamont, A. and I. Cross (1994). "Children's Cognitive Representations of Musical Pitch." Music Perception: An Interdisciplinary Journal **12**(1): 27-55.

Laporte, M. N. and K. Zuberbühler (2010). "Vocal greeting behaviour in wild chimpanzee females." Animal Behaviour **80**(3): 467-473.

Large, E. W. (1994). "Dynamic representation of musical structure." The Ohio State University Press.

Large, E. W. (2008). "Resonating to musical rhythm: theory and experiment." The psychology of time: 189-232.

Large, E. W. (2010). Neurodynamics of music. Music perception, Springer**:** 201-231.

Large, E. W. and P. M. Gray (2015). "Spontaneous tempo and rhythmic entrainment in a bonobo (Pan paniscus)." Journal of Comparative Psychology **129**(4): 317.

Large, E. W., J. A. Herrera, et al. (2015). "Neural networks for beat perception in musical rhythm." Frontiers in systems neuroscience **9**.

Large, E. W. and M. R. Jones (1999). "The dynamics of attending: How people track time-varying events." Psychological review **106**(1): 119.

Large, E. W. and J. F. Kolen (1994). "Resonance and the perception of musical meter." Connection science **6**(2-3): 177-208.

Large, E. W. and J. S. Snyder (2009). "Pulse and meter as neural resonance." Annals of the New York Academy of sciences **1169**(1): 46-57.

Larson, C., H. Liu, et al. (2008). Laryngeal muscle responses to voice auditory feedback perturbations. International Conference on Voice Physiology and Biomechanics, Tampere, Finland.

Larson, C. R. (1998). "Cross-modality influences in speech motor control: the use of pitch shifting for the study of F0 control." Journal of communication disorders **31**(6): 489-503.

Larson, C. R. and D. A. Robin (2016). "Sensory Processing: Advances in Understanding Structure and Function of Pitch-Shifted Auditory Feedback in Voice Control." AIMS Neuroscience **3**(1): 22-39.

Larson, C. R., J. Sun, et al. (2007). "Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F 0 and amplitude control a." The Journal of the Acoustical Society of America **121**(5): 2862-2872.

Laukka, P. (2004). Vocal expression of emotion: Discrete-emotions and dimensional accounts. Faculty of Social Sciences, University of Uppsala. **PhD**.

Laukka, P., T. Eerola, et al. (2013). "Universal and culture-specific factors in the recognition and performance of musical affect expressions." Emotion **13**(3): 434.

Laukka, P., P. Juslin, et al. (2005). "A dimensional approach to vocal expression of emotion." Cognition and Emotion **19**(5): 633-653.

Launay, J., B. Tarr, et al. (2016). "Synchrony as an Adaptive Mechanism for Large-Scale Human Social Bonding." Ethology **122**(10): 779-789.

Lavy, M. (2001). Emotion and the experience of listening to music: A framework for empirical research. Centre for Music and Science, University of Cambridge. **PhD**.

Lawson, G. and F. D'Errico (2002). Microscopic, experimental and theoretical re-assessment of upper Paleolithic bird-bone pipes from Isuritz, France: Ergonomics of design, systems of notation and the origins of musical traditions. Studien zur Musikachaelogie III. E. Hickmann, A. Kilmer and R. Eichman. Rahden, Verlag Marie Leidorf.

Lazarus, R. (1982). "Thoughts on the relations between emotion and cognition." American Psychologist **37**: 1019-1034.

League, P. (2016). Matters of taste and time in Anatolian Greek music. Time in Music. Conference, University of Oxford.

Leakey, R. (1994). The origin of humankind. New York, Basic Books.

LeDoux, J. (1996). The emotional brain. New York, Simon & Schuster.

Lehmann, A., E. Skoe, et al. (2015). "Impairments in musical abilities reflected in the auditory brainstem: evidence from congenital amusia." European Journal of Neuroscience **42**(1): 1644-1650.

Lehmann, J., A. Korstjens, et al. (2007). "Group size, grooming and social cohesion in primates." Animal Behaviour **74**(6): 1617-1629.

Lejeune, H. and J. Wearden (2006). "Scalar properties in animal timing: Conformity and violations." The Quarterly journal of experimental psychology **59**(11): 1875-1908.

Lemasson, A., L. Glas, et al. (2011). "Youngsters do not pay attention to conversational rules: is this so for nonhuman primates?" Scientific reports **1**: 22.

Lerdahl, F. and R. Jackendorff (1983). A generative theory of tonal music. Cambridge, Massachusetts, MIT Press.

Lester, B. M., J. Hoffman, et al. (1985). "The rhythmic structure of mother-infant interaction in term and preterm infants." Child development: 15-27.

Levi, E. (1994). Music in the Third Reich. New York, St. Martin's Press.

Levinson, S. C. (2006). On the human" interaction engine". Wenner-Gren Foundation for Anthropological Research, Symposium 134, Berg.

Levinson, S. C. (2011). Cross–cultural universals and communication structures. the Ernst Strungmann Forum: Language, Music and the Brain.

Levinson, S. C. (2016). "Turn-taking in human communication–origins and implications for language processing." Trends in cognitive sciences **20**(1): 6-14.

Levinson, S. C. and F. Torreira (2015). "Timing in turn-taking and its implications for processing models of language." Frontiers in psychology **6**.

Levman, B. (2000). "Western theories of music origin, historical and modern." Musicae Scientiae **4**: 185-211.

Lewin, R. (2009). Human evolution: an illustrated introduction, John Wiley & Sons.

Lewis, J. (2009). "13 As well as words: Congo Pygmy hunting, mimicry, and play." The cradle of language **12**: 236.

Lewis, J. (2013). A cross-cultural perspective on the significance of music and dance to culture and society insight from BaYaka pygmies, MIT Press.

Lewis, J. D. (2014). Pygmy hunter-gatherer egalitarian social organization: the case of the Mbendjele BaYaka, Transaction Publishers.

Lewis, P. A., A. Wing, et al. (2004). "Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping." Neuropsychologia **42**(10): 1301-1312.

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

Liberman, A. M. (1996). Speech: A special code, MIT press.

Liberman, M. (1984). "Intonational invariance under changes in pitch range and length." Language sound structure.

Lieberman, P. (1984). The biology and evolution of language. Cambridge, MA, MIT Press.

Lieberman, P. (2017). "Comment on “Monkey vocal tracts are speech-ready”." Science Advances **3**(7): e1700442.

Lieberman, P., E. S. Crelin, et al. (1972). "Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee." American Anthropologist **74**(3): 287-307.

Lieberman, P. and R. C. McCarthy (2013). The evolution of speech and language. Handbook of paleoanthropology, Springer**:** 1-41.

Lieberman, P., M. Sawashima, et al. (1970). "The articulatory implementation of the breath-group and prominence: crico-thyroid muscular activity in intonation." Language: 312-327.

Lim, K.-B., Y.-K. Kim, et al. (2013). "The therapeutic effect of neurologic music therapy and speech language therapy in post-stroke aphasic patients." Annals of rehabilitation medicine **37**(4): 556-562.

Lindblom, B. and J. Sundberg (2007). The human voice in speech and singing. Springer handbook of acoustics, Springer**:** 669-712.

List, G. (1963). The Boundaries of Speech and Song. Ethnomusicology, 7(1), 1-16. Doi 10.2307/924141.

Liszkowski, U., M. Carpenter, et al. (2004). "Twelve-month-olds point to share attention and interest." Developmental science **7**(3): 297-307.

Liszkowski, U., M. Carpenter, et al. (2006). "12-and 18-month-olds point to provide information for others." Journal of Cognition and Development **7**(2): 173-187.

Liu, F., A. H. Chan, et al. (2016). "Pitch perception and production in congenital amusia: Evidence from Cantonese speakers." The Journal of the Acoustical Society of America **140**(1): 563-575.

Liu, F., A. R. Maggu, et al. (2015). "Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers." Frontiers in human neuroscience **8**: 1029.

Liu, F., A. D. Patel, et al. (2010). "Intonation processing in congenital amusia: discrimination, identification and imitation." Brain **133**(6): 1682-1693.

Liu, F., Y. Xu, et al. (2012). "Differential recognition of pitch patterns in discrete and gliding stimuli in congenital amusia: Evidence from Mandarin speakers." Brain and cognition **79**(3): 209-215.

Liu, H., R. Behroozmand, et al. (2011). "Laryngeal electromyographic responses to perturbations in voice pitch auditory feedback." The Journal of the Acoustical Society of America **129**(6): 3946-3954.

Liu, H. and C. R. Larson (2007). "Effects of perturbation magnitude and voice F 0 level on the pitch-shift reflex." The Journal of the Acoustical Society of America **122**(6): 3671-3677.

Liu, H., Y. Xu, et al. (2009). "Attenuation of vocal responses to pitch perturbations during Mandarin speech." The Journal of the Acoustical Society of America **125**(4): 2299-2306.

Livingstone, F. (1973). "Did the Austalopithecines sing?" Current Anthropology **14**: 25-29.

Livingstone, S. R. and W. F. Thompson (2009). "The emergence of music from the Theory of Mind." Musicae Scientiae **13**(2_suppl): 83-115.

Locke, J. (1996). "Why do infants begin to talk? Language as an unintended consequence." Journal of Child Language **23**: 251-268.

Loehr, J. D., E. W. Large, et al. (2011). "Temporal coordination and adaptation to rate change in music performance." Journal of experimental psychology: Human perception and performance **37**(4): 1292.

Loehr, J. D. and C. Palmer (2009). "Subdividing the beat: Auditory and motor contributions to synchronization." Music Perception: An Interdisciplinary Journal **26**(5): 415-425.

Lomax, A. (1968). Folk song style and culture. New Brunswick, Transaction Books.

Lombardo, M. V. and S. Baron-Cohen (2011). "The role of the self in mindblindness in autism." Consciousness and cognition **20**(1): 130-140.

London, J. (2004/12). Hearing in time: Psychological aspects of musical metre. Oxford, Oxford University Press.

London, J., B. Burger, et al. (2016). "Speed on the dance floor: Auditory and visual cues for musical tempo." Acta psychologica **164**: 70-80.

Lorenz, K. (1950). "The comparative method in studying innate behaviour patterns." Physiological mechanisms in animal behaviour(17).

Lorenz, K. Z. (1966). "Evolution of ritualization in the biological and cultural spheres." Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **251**(772): 273-284.

Lotze, M., M. Wittmann, et al. (1999). "Daily rhythm of temporal resolution in the auditory system." Cortex **35**(1): 89-100.

Loui, P. (2016). "Absolute Pitch." The Oxford handbook of music psychology: 81.

Loui, P., D. L. Wessel, et al. (2010). "Humans rapidly learn grammatical structure in a new musical scale." Music Perception: An Interdisciplinary Journal **27**(5): 377-388.

Lucas, G., M. Clayton, et al. (2011). "Inter-group entrainment in Afro-Brazilian Congado ritual." Empirical musicology review. **6**(2): 75-102.

Luft, A. R., S. McCombe-Waller, et al. (2004). "Repetitive bilateral arm training and motor cortex activation in chronic stroke: a randomized controlled trial." Jama **292**(15): 1853-1861.

Lumsden, J., L. K. Miles, et al. (2014). "Sync or sink? Interpersonal synchrony impacts self-esteem." Frontiers in psychology **5**: 1064.

Lumsden, J., L. K. Miles, et al. (2012). "Who syncs? Social motives and interpersonal coordination." Journal of Experimental Social Psychology **48**(3): 746-751.

Lundqvist-Persson, C. (2001). "Correlation between level of self-regulation in the newborn infant and developmental status at two years of age." Acta Paediatrica **90**(3): 345-350.

Lynch, M., R. Eilers, et al. (1990). "Innateness, experience, and music perception." Psychological Science **1**: 272-276.

Machin, A. J. and R. I. Dunbar (2011). "The brain opioid theory of social attachment: a review of the evidence." Behaviour **148**(9-10): 985-1025.

Maclarnon, A. (1993). The vertebral canal. The Nariokotome *Homo erectus* skeleton. A. Walker and R. Leakey. Cambridge, MA, Harvard University Press**:** 359-390.

Maclarnon, A. and G. Hewitt (1999). "THe evolution of human speech: the role of enhanced breathing control." American Journal of Physical Anthropology **109**: 341-363.

Maclarnon, A. and G. Hewitt (2004). "Increased breathing control: Another factor in the evolution of human language." Evolutionary Anthropology **13**(5): 181-197.

Macrae, C. N., O. K. Duffy, et al. (2008). "A case of hand waving: Action synchrony and person perception." Cognition **109**(1): 152-156.

Madison, G. (1992). Drumming performance with and without clicktrack-the validity of the internal clock in expert synchronisation. Proc. Fourth Workshop on Rhythm Perception & Production.

Madison, G. (2000). On the nature of variability in isochronous serial interval production. Rhythm: Perception and Production. P. W. Desain, L. , Swets & Zeitlinger Publishers.

Madison, G. (2001). "Variability in isochronous tapping: Higher order dependencies as a function of intertap interval." Journal of Experimental Psychology Human Perception and Performance **27**(2): 411-422.

Maestripieri, D. and J. Call (1996). Mother-infant communication in primates. Advances in the Study of Behavior, Elsevier. **25:** 613-642.

Magriples, U. and J. Laitman (1987). "Developmental change in the position of the fetal human larynx." American Journal of Physical Anthropology **72**: 463-472.

Magyari, L., M. C. Bastiaansen, et al. (2014). "Early anticipation lies behind the speed of response in conversation." Journal of cognitive neuroscience.

Magyari, L. and J. P. de Ruiter (2012). "Prediction of turn-ends based on anticipation of upcoming words." Frontiers in psychology **3**.

Maher, T. F. (1976). "" Need for Resolution" Ratings for Harmonic Musical Intervals: A Comparison between Indians and Canadians." Journal of Cross-Cultural Psychology **7**(3): 259-276.

Maiese, M. (2015). "Giovanna Colombetti, The feeling body: affective science meets the enactive mind, MIT Press, 2013, 288pp, Hardcover, $40.00, ISBN: 9780262019958." Phenomenology and the Cognitive Sciences **14**(4): 973-978.

Maiese, M. (2017). "Can the mind be embodied, enactive, affective, and extended?" Phenomenology and the Cognitive Sciences: 1-19.

Mainka, S., R. Spintge, et al. (2016). Music Therapy in Medical and Neurological Rehabilitation Settings. In Hallam, S., Cross, I. & Thaut, M. The Oxford handbook of music psychology 857 – 877. Oxford University Press.

Malcolm, M. P., C. Massie, et al. (2009). "Rhythmic auditory-motor entrainment improves hemiparetic arm kinematics during reaching movements: a pilot study." Topics in stroke rehabilitation **16**(1): 69-79.

Malloch, S. (1999/2000). "Mothers and infants and communicative musicality." Musicae Scientiae, special issue: Rhythm, musical narrative, and the origins of human communcation: 29-58.

Malloch, S. and C. Trevarthen (2009). Communicative musicality: Exploring the basis of human companionship, Oxford University Press, USA.

Malloch, S. and C. Trevarthen (2009). "Musicality: Communicating the vitality and interests of life." Communicative musicality: Exploring the basis of human companionship **1**: 1-10.

Mann, N., L. Marshall-Ball, et al. (2003). "The complex song duet of the plain wren." Condor **105**: 672-682.

Manstead, A., N. Frijda, et al., Eds. (2004). Feeling and emotions: The Amsterdam symposium. Cambridge, Cambridge University Press.

Marler, P. (1970). "Birdsong and speech development: Could there be parallels? There may be basic rules governing vocal learning to which many species conform, including man." American scientist **58**(6): 669-673.

Martin-Vivaldi, M., J. Palomino, et al. (1999). "Function of song in the Hoopoe Upupa epops " Bird Study **46**(1): 104-111.

Martin, A. and L. R. Santos (2014). "The origins of belief representation: Monkeys fail to automatically represent others' beliefs." Cognition **130**(3): 300-308.

Martinelli, D. (2010). A Critical Companion to Zoosemiotics:: People, Paths, Ideas, Springer Science & Business Media.

Martínez, I., J. L. Arsuaga, et al. (2008). "Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain)." Journal of Human Evolution **54**(1): 118-124.

Martinez, I., A. Rosas, et al. (2004). "Auditory capacities in middle pleistocence humans from the Sierra de Atapuerca in Spain." Proceedings of the National Academy of Sciences **101**(27): 9976-9981.

Mas-Herrero, E., R. J. Zatorre, et al. (2014). "Dissociation between musical and monetary reward responses in specific musical anhedonia." Current Biology **24**(6): 699-704.

Masali, M. and M. Cremasco (2006). "Hoc alterum auditus organi ossiculum est: Ear Ossicles in Physical Anthropology " human Evolution **21**(1): 1-17.

Masataka, N. (2006). "Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents." Developmental science **9**(1): 46-50.

Masataka, N. (2008). Implication of the human musical faculty for evolution of language. The Origins of Language, Springer**:** 133-151.

Masataka, N. (2009). "The origins of language and the evolution of music: A comparative perspective." Physics of Life Reviews **6**(1): 11-22.

Masataka, N. and L. Perlovsky (2013). "Cognitive interference can be mitigated by consonant music and facilitated by dissonant music." Scientific reports **3**.

Maslow, A. (1968). Toward a psychology of being (2nd edition). New York, Van Nostrand Reinhold.

Maslow, A. H. (2013). Toward a psychology of being, Simon and Schuster.

Massie, C., M. P. Malcolm, et al. (2009). "The effects of constraint-induced therapy on kinematic outcomes and compensatory movement patterns: an exploratory study." Archives of physical medicine and rehabilitation **90**(4): 571-579.

Mates, J. (1994). "A model of synchronization of motor acts to a stimulus sequence." Biological cybernetics **70**(5): 463-473.

Mates, J. and G. Aschersleben (2000). "Sensorimotor synchronization: the impact of temporally displaced auditory feedback." Acta psychologica **104**(1): 29-44.

Mates, J., U. Müller, et al. (1994). "Temporal integration in sensorimotor synchronization." Journal of cognitive neuroscience **6**(4): 332-340.

Mates, J., T. Radil, et al. (1992). "Cooperative tapping: Time control under different feedback conditions." Attention, Perception, & Psychophysics **52**(6): 691-704.

May, J. and P. Hamilton (1980). "Effects of musically evoked affect on women's interpersonal attraction toward and perceptual judgments of physical attractiveness of men." Motivation and Emotion **4**: 217-228.

Mayville, J. M., K. J. Jantzen, et al. (2002). "Cortical and subcortical networks underlying syncopated and synchronized coordination revealed using fMRI." Human brain mapping **17**(4): 214-229.

McAllester, D. (1954). Enemy Way Music. Peabody Museum Papers 41 (3), Cambridge, Mass. Harvard University Press.

McAllester, D. P. (1954). "Enemy way music." Cambridge, Mass.: Peabody Museum, Harvard University.

McAllester, D. P. (1971). "Some thoughts on" universals" in world music." Ethnomusicology **15**(3): 379-380.

McAngus Todd, N. P. (1992). "The dynamics of dynamics: A model of musical expression." The Journal of the Acoustical Society of America **91**(6): 3540-3550.

McAngus Todd, N. P., D. J. O'Boyle, et al. (1999). "A sensory-motor theory of rhythm, time perception and beat induction." Journal of New Music Research **28**(1): 5-28.

McAuley, J. D. (1995). "Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing."

McAuley, J. D., D. Frater, et al. (2006). Detecting changes in timing: Evidence for two modes of listening. The Proceedings of the 9th International Conference on Music Perception and Cognition.

McAuley, J. D., M. R. Jones, et al. (2006). "The time of our lives: life span development of timing and event tracking." Journal of Experimental Psychology: General **135**(3): 348.

McClintock, M. K. (1984). "Group mating in the domestic rat as a context for sexual selection: Consequences for the analysis of sexual behavior and neuroendocrine responses." Advances in the Study of Behavior **14**: 1-50.

McConnell, P. (1991). Lessons from animal trainors: the effect of acoustic structure on an animal's response. Perspectives in ethology, vol. 9: Human understanding and animal awareness. P. Bateson and P. Klopfer. New York, Plenum Press**:** 165-187.

McDermott, J. and M. Hauser (2004). "Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate." Cognition **94**(2): B11-B21.

McDermott, J. and M. Hauser (2005). "The origins of music: Innateness, uniqueness and evolution." Music Perception **23**(1): 29-60.

McDermott, J. H. (2009). "What can experiments reveal about the origins of music?" Current Directions in Psychological Science **18**(3): 164-168.

McDermott, J. H., A. J. Lehr, et al. (2010). "Individual differences reveal the basis of consonance." Current Biology **20**(11): 1035-1041.

McDermott, J. H. and A. J. Oxenham (2008). "Music perception, pitch, and the auditory system." Current opinion in neurobiology **18**(4): 452-463.

McDermott, J. H., A. F. Schultz, et al. (2016). "Indifference to dissonance in native Amazonians reveals cultural variation in music perception." Nature **535**(7613): 547-550.

McGrew, W. (1998). "Culture in nonhuman primates?" Annual Review of Anthropology **27**: 301-328.

McKinney, M., M. Tramo, et al. (2001). "Neural correlates of musical dissonance in the inferior colliculus." Physiological and psychophysical bases of auditory function (Breebaart DJ, Houtsma AJM, Kohlrausch A, Prijs VF, Schoonhoven R, eds): 83-89.

McLachlan, N., D. Marco, et al. (2013). "Consonance and pitch." Journal of Experimental Psychology: General **142**(4): 1142.

McNeill, W. (1995). Keeping together in time: Dance and drill in human history. Cambridge, MA, Harvard University Press.

Meck, W. H. (2003). Functional and neural mechanisms of interval timing, CRC Press.

Meck, W. H. (2005). "Neuropsychology of timing and time perception." Brain and cognition **58**(1): 1-8.

Mega, M., J. Cummings, et al. (1997). The limbic system: an anatomic, phylogenetic, and clinical perspective. The neuropsychiatry of limbic and subcortical disorders. S. Salloway, P. Malloy and J. Cummings. Washington, D.C., American Psychiatric Press**:** 3-18.

Mehr, S. A. and M. M. Krasnow (2017). "Parent-offspring conflict and the evolution of infant-directed song." Evolution and Human Behavior.

Mellars, P. (2006). "A new radiocarbon revolution and the dispersal of modern humans in Eurasia." Nature **439**(7079): 931-935.

Meltzoff, A. N. (1985). "The roots of social and cognitive development: Models of man's original nature." Social perception in infants: 1-30.

Meltzoff, A. N. (1990). Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling, and self practice in infancy. The John D. and Catherine T. MacArthur foundation series on mental health and development. The self in transition: Infancy to childhood. D. C. M. Beeghly. Chicago, University of Chicago Press**:** 139-164.

Meltzoff, A. N. (1993). The role of imitation in understanding persons and developing a theory of mind. Understanding other minds: Perspectives from autism. S. Baron-Cohen and H. T.-F. D. Cohen. New York, Oxford University Press**:** 335–366.

Meltzoff, A. N. (2007). "'Like me': a foundation for social cognition." Developmental science **10**(1): 126-134.

Meltzoff, A. N. and M. K. Moore (1989). "Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms." Developmental Psychology **25**(6): 954.

Mendelson, M. (1986). "Perception of the temporal pattern of motion in infancy." Infant Behaviour and Development **9**: 231-243.

Méndez-Cárdenas, M. G. and E. Zimmermann (2009). "Duetting—A mechanism to strengthen pair bonds in a dispersed pair-living primate (Lepilemur edwardsi)?" American Journal of Physical Anthropology **139**(4): 523-532.

Mendez, J. C., L. Prado, et al. (2011). "Temporal and spatial categorization in human and non-human primates." Frontiers in Integrative Neuroscience **5**.

Mendez, M. (2001). "Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke." Cortex **37**: 139-150.

Mendl, M., O. H. Burman, et al. (2010). "An integrative and functional framework for the study of animal emotion and mood." Proceedings of the Royal Society of London B: Biological Sciences **277**(1696): 2895-2904.

Mendoza, G. and H. Merchant (2014). "Motor system evolution and the emergence of high cognitive functions." Progress in neurobiology **122**: 73-93.

Menon, V. and D. J. Levitin (2005). "The rewards of music listening: response and physiological connectivity of the mesolimbic system." Neuroimage **28**(1): 175-184.

Merchant, H. and R. Bartolo (2017). "Primate beta oscillations and rhythmic behaviors." Journal of Neural Transmission: 1-10.

Merchant, H., R. Bartolo, et al. (2014). Neurophysiology of timing in the hundreds of milliseconds: multiple layers of neuronal clocks in the medial premotor areas. Neurobiology of Interval Timing, Springer**:** 143-154.

Merchant, H., J. Grahn, et al. (2015). "Finding the beat: a neural perspective across humans and non-human primates." Phil. Trans. R. Soc. B **370**(1664): 20140093.

Merchant, H. and H. Honing (2013). "Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis." Frontiers in neuroscience **7**.

Merchant, H., O. Pérez, et al. (2013). "Interval tuning in the primate medial premotor cortex as a general timing mechanism." Journal of Neuroscience **33**(21): 9082-9096.

Merker, B. (2000). Synchronous chorusing and human origins. The origins of music. N. Wallin, B. Merker and S. Brown. Cambridge MA, MIT Press**:** 315-327.

Merker, B. (2006). "The uneven interface between culture and biology in human music." Music Perception: An Interdisciplinary Journal **24**(1): 95-98.

Merker, B. (2009). From ape culture to ritual culture and language via vocal learning for song: THe singular path to human uniqueness. Communicative musicality. S. Malloch and C. Trevarthan. Oxford, Oxford University Press.

Merker, B. H., G. S. Madison, et al. (2009). "On the role and origin of isochrony in human rhythmic entrainment." Cortex **45**(1): 4-17.

Merriam, A. (1964). The anthropology of music. Evanston, Ill., NorthWestern University Press.

Merriam, A. (1967). Ethnomusicology of the Flathead Indians. Chicago, Aldine Press.

Merriam, A. P. (1967). Music and the Origin of the Flathead Indians: A Problem in Culture History, Indiana University Research Center in Anthropology, Folklore and Linguistics.

Metz-Lutz, M. and E. Dahl (1984). "Analysis of word comprehension in a case of pure word deafness." Brain and Language **23**(1): 13-25.

Meyer, L. (1956). Emotion and meaning in music. London, University of Chicago Press.

Miall, C. (2010). "Motor control: correcting errors and learning from mistakes." Current Biology **20**(14): R596-R598.

Michael, J., N. Sebanz, et al. (2016). "Observing joint action: Coordination creates commitment." Cognition **157**: 106-113.

Micheyl, C., C. M. Ryan, et al. (2012). "Further evidence that fundamental-frequency difference limens measure pitch discrimination." The Journal of the Acoustical Society of America **131**(5): 3989-4001.

Michon, J. A. (1967). Timing in temporal tracking, Institute for Perception RVO-TNO Soesterberg, The Netherlands.

Miell, D., R. MacDonald, et al., Eds. (2005). Musical Communication. Oxford, Oxford University Press.

Miles, L. K., J. L. Griffiths, et al. (2010). "Too late to coordinate: Contextual influences on behavioral synchrony." European Journal of Social Psychology **40**(1): 52-60.

Miles, L. K., L. K. Nind, et al. (2010). "Moving memories: Behavioral synchrony and memory for self and others." Journal of Experimental Social Psychology **46**(2): 457-460.

Miles, L. K., L. K. Nind, et al. (2009). "The rhythm of rapport: Interpersonal synchrony and social perception." Journal of Experimental Social Psychology **45**(3): 585-589.

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

Miller, L. (1989). Musical savants. Exceptional skill in the mentally retarded. Hillsdale, NJ, Erlbaum.

Mills, L. and G. B. Rollman (1980). "Hemispheric asymmetry for auditory perception of temporal order." Neuropsychologia **18**(1): 41-47.

Mitani, J. (1990). "Experimental field studies of Asian ape social systems." International Journal of Primatology **11**(2): 103-126.

Mitani, J., K. Hunley, et al. (1999). "Geographic variation in the calls of wild chimpanzees: a reassessment." American Journal of primatology **47**(2): 133-151.

Mithen, S. (1996). "The prehistory of the mind: a search for the origins of art, science and religion." London and New York: Thames and Hudson.

Mithen, S. (2005). The singing neanderthals: The origins of music, language, mind and body. London, Weidenfeld and Nicolson.

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

Mizuno, Y. and H. Takashita (2002). "Behavioral development of chimpanzees in the first month of life: Observation of mother-infant pairs at night." Japanese Psychological Review **45**: 352-364.

Mogan, R., R. Fischer, et al. (2017). "To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect." Journal of Experimental Social Psychology **72**: 13-20.

Molinari, M., M. Leggio, et al. (2003). "Neurobiology of rhythmic motor entrainment." Annals of the New York Academy of Sciences **999**: 313-321.

Moll, H. and A. N. Meltzoff (2011). "Foundation in Joint Attention." Perception, causation, and objectivity **286**.

Moll, H. and M. Tomasello (2004). "12-and 18-month-old infants follow gaze to spaces behind barriers." Developmental science **7**(1).

Moore, B. C. (1993). "Temporal analysis in normal and impaired hearing." Annals of the New York Academy of sciences **682**(1): 119-136.

Moore, B. C. and B. R. Glasberg (1990). "Frequency discrimination of complex tones with overlapping and non-overlapping harmonics." The Journal of the Acoustical Society of America **87**(5): 2163-2177.

Moore, G. A. and S. D. Calkins (2004). "Infants' vagal regulation in the still-face paradigm is related to dyadic coordination of mother-infant interaction." Developmental Psychology **40**(6): 1068.

Moore, G. A., A. L. Hill-Soderlund, et al. (2009). "Mother–infant vagal regulation in the face-to-face still-face paradigm is moderated by maternal sensitivity." Child development **80**(1): 209-223.

Moore, R. E., J. Estis, et al. (2008). "Pitch discrimination and pitch matching abilities with vocal and nonvocal stimuli." Journal of Voice **22**(4): 399-407.

Moore, R. E., C. Keaton, et al. (2007). "The role of pitch memory in pitch discrimination and pitch matching." Journal of Voice **21**(5): 560-567.

Moreau, P., P. Jolicœur, et al. (2013). "Pitch discrimination without awareness in congenital amusia: evidence from event-related potentials." Brain and cognition **81**(3): 337-344.

Moreno, S., E. Bialystok, et al. (2011). "Short-term music training enhances verbal intelligence and executive function." Psychological science **22**(11): 1425-1433.

Morgan, E. (1995). The descent of the child: Human evolution from a new perspective. New York, Oxford University Press.

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

Morley, I. (2003). The evolutionary origins and archaeology of music. An investigation into the prehistory of human musical capacities and behaviours. Department of Archaeology, University of Cambridge.

Morley, I. (2009). "Ritual and music: parallels and practice and the Paleolithic." Becoming human: innovation in prehistoric material and spiritual culture: 159-175.

Morley, I. (2013). The prehistory of music: human evolution, archaeology, and the origins of musicality, Oxford University Press.

Morrongiello, B., S. Trehub, et al. (1985). "Children's perception of melodies: The role of contour, frequency, and rate of presentation." Journal of Experimental Child Psychology **40**: 279-292.

Morton, E. (1994). "Sound symbolism and its role in non-human vertebrate communication." Acta Psychologia **20**: 308-319.

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.

Moynihan, M. (1956). "Notes on the behavior of some North American gulls. I. Aerial hostile behavior." Behaviour **10**(1): 126-177.

Muente, T., E. Altenmueller, et al. (2002). "The musician's brain as a model of neuroplasticity." Nature Reviews **3**: 473-478.

Müller, A. E. and G. Anzenberger (2002). "Duetting in the titi monkey Callicebus cupreus: structure, pair specificity and development of duets." Folia Primatologica **73**(2-3): 104-115.

Müller, K., G. Aschersleben, et al. (2001). "Modalityspecific central control units in sensorimotor synchronization." Manuscript submitted for publication.

Muruyama, J., T. Kashiwagi, et al. (2004). "Impaired pitch production and preserved rhythm production in a right brain-damaged patient with amusia." Brain and Cognition **56**: 36-42.

Näätänen, R. and C. Escera (2000). "Mismatch negativity." Audiology & Neuro-otology **5**(3): 105-110.

Nadel, S. and T. Baker (1930). "The origins of music." The Musical Quarterly **16**(4): 531-546.

Nagasaki, H. (1990). "Rhythm and variability of timing in periodic tapping." Human Movement Science **9**(2): 177-194.

Nakahara, H., S. Furuya, et al. (2009). "Emotion-related Changes in Heart Rate and Its Variability during Performance and Perception of Music." Annals of the New York Academy of sciences **1169**(1): 359-362.

Nakamura, J. and M. Csikszentmihalyi (2014). The concept of flow. Flow and the foundations of positive psychology, Springer**:** 239-263.

Nan, Y., T. Knosche, et al. (2006). "The perception of musical phrase structure: a cross-cultural ERP study." Brain Research **1094**(1): 179-191.

Nan, Y., Y. Sun, et al. (2010). "Congenital amusia in speakers of a tone language: association with lexical tone agnosia." Brain **133**(9): 2635-2642.

Narmour, E. (1990). The analysis and cognition of basic melodic structures. Chicago, University of Chicago Press.

Natke, U., T. M. Donath, et al. (2003). "Control of voice fundamental frequency in speaking versus singing." The Journal of the Acoustical Society of America **113**(3): 1587-1593.

Natke, U. and K. T. Kalveram (2001). "Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables." Journal of Speech, Language, and Hearing Research **44**(3): 577-584.

Negus, K. (1997). Popular music in theory: An introduction. Hanover, NH, University Press of New England.

Nelson, S. (2002). Melodic improvisation on a twelve-bar blues: an investigation of physical and historical aspects, and their contribution to performance. London, City University London. **PhD**.

Nespor, M., M. Shukla, et al. (2011). "49 Stress-timed vs. Syllable-timed Languages." In M. van Oostendorp, C. J. Ewen, E. Hume, K. Rice (eds.) The Blackwell Companion to Phonology. Chichester, Uk: Wiley-Blackwell, pp. 1147-1159.

Nettl, B. (1983). The study of ethnomusicology: Twenty-nine issues and concepts, University of Illinois Press.

Nettl, B. (2000). An ethnomusicologist contemplates musical universals. The origins of music. N. Wallin, B. Merker and S. Brown. Cambridge, MA, MIT Press.

Nettl, B. (2005/2010). The study of ethnomusicology: thirty-three discussions (2nd/3rd Editions). University of Illinois Press.

Nettl, B. (2015). The study of ethnomusicology: thirty-three discussions (4th Edition). University of Illinois Press.

Nettl, B. (2016). Excursions in world music, Taylor & Francis.

Neumann, O. (1990). "Direct parameter specification and the concept of perception." Psychological Research **52**(2-3): 207-215.

Newman, J. (2003). "Auditory communication and central auditory mechanisms in the squirrel monkeys: past and present." Primate Audition: Ethology and Neurobiology **13**: 198-203.

Newman, J. and P. Goedeking (1992). Noncategorical vocal communication in primates: the example of common marmoset phee calls. Nonverbal vocal communication: comparative and developmental approaches. H. Papousek, U. Jurgens and M. Papousek. Cambridge, Cambridge University Press**:** 87-101.

Nguyen, S., B. Tillmann, et al. (2009). "Tonal language processing in congenital amusia." Annals of the New York Academy of Sciences **1169**(1): 490-493.

Nielzen, S. and Z. Cesarec (1981). "On the perception of music as a function of musical structure." Psychology of Music **10**: 7-17.

Nishimura, T. (2006). "Descent of the larynx in chimpanzees: Mosaic and multiple-step evolution of the foundations for human speech." Matsuzawa T, Tomonaga M, Tanaka M, edsitors. Cognitive development in chimpanzees. Tokyo: 75-95.

Nishimura, T., T. Oishi, et al. (2008). "Development of the supralaryngeal vocal tract in Japanese macaques: implications for the evolution of the descent of the larynx." American journal of physical anthropology **135**(2): 182-194.

Nolan, F. and H.-S. Jeon (2014). "Speech rhythm: a metaphor?" Phil. Trans. R. Soc. B **369**(1658): 20130396.

Norman-Haignere, S., N. Kanwisher, et al. (2013). "Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex." Journal of Neuroscience **33**(50): 19451-19469.

Norman-Haignere, S. V., P. Albouy, et al. (2016). "Pitch-responsive cortical regions in congenital amusia." Journal of Neuroscience **36**(10): 2986-2994.

Norman, D. and T. Shallice (1980). Attention to action: Willed and automatic control of behaviour. Reprinted in M. Gazzaniga (ed)(2000) Cognitive Neuroscience: A Reader, Blackwell.

Norman, D. A. and T. Shallice (1986). Attention to action. Consciousness and self-regulation, Springer**:** 1-18.

North, A. and D. Hargreaves (1997). Music and consumer behaviour. The social psychology of music. D. Hargreaves and A. North. Oxford, Oxford University Press**:** 268-289.

North, A., M. Tarrant, et al. (2004). "The effects of music on helping behaviour: A field study." Environment and Behaviour **36**: 266-275.

Nöth, W. (1995). Handbook of semiotics, Indiana University Press.

Nowicki, L., W. Prinz, et al. (2013). "Mutual adaptive timing in interpersonal action coordination." Psychomusicology: Music, Mind, and Brain **23**(1): 6.

Nummela, S. (1995). "Scaling of the mammalian middle ear." Hearing Research **85**(1-2): 18-30.

Nyklicek, I., J. Thayer, et al. (1997). "Cardiorespiratory differentiation of musically-induced emotions." Journal of Psychophysiology **11**: 304-321.

Oatley, K. (1992). Best laid schemes. The psychology of emotions. Cambridge, MA, Harvard University Press.

Oatley, K. and J. Jenkins (1996). Understanding emotions. Oxford, UK, Blackwell.

Ogden, R. and S. Hawkins (2015). Entrainment as a basis for co-ordinated actions in speech. International Congress of the Phonetic Sciences, York.

Ohala, J. (1984). "An ethological perspective on common cross-language utilization of F0 in voice." Phonetica **41**: 1-16.

Okada, B. M., L. Lachs, et al. (2012). "Interpreting tone of voice: musical pitch relationships convey agreement in dyadic conversation." The Journal of the Acoustical Society of America **132**(3): 208-214.

Okamoto, S., M. Tomonaga, et al. (2002). "An infant chimpanzee (Pan troglodytes) follows human gaze." Animal cognition **5**(2): 107-114.

Okanoya, K. and B. Merker (2007). Neural substrates for string-context mutual segmentation: A path to human language. Emergence of communication and language, Springer**:** 421-434.

Oller, D. K. and U. Griebel (2008). Evolution of communicative flexibility: complexity, creativity, and adaptability in human and animal communication, MIT Press (MA).

Olsho, L., C. Schoon, et al. (1982). "Preliminary data on frequency discrimination in infancy." Journal of the Acoustical Society of America **71**: 509-511.

Olson, D. R. and D. Kamawar (1999). "The theory of ascriptions." In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), Developing theories of intention: Social understanding and self-control (pp. 153-166). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Omigie, D., D. Müllensiefen, et al. (2012). "The experience of music in congenital amusia." Music Perception: An Interdisciplinary Journal **30**(1): 1-18.

Omigie, D., M. T. Pearce, et al. (2013). "Electrophysiological correlates of melodic processing in congenital amusia." Neuropsychologia **51**(9): 1749-1762.

Onishi, K. H. and R. Baillargeon (2005). "Do 15-month-old infants understand false beliefs?" Science **308**(5719): 255-258.

Opitz, B., T. Rinne, et al. (2002). "Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results." Neuroimage **15**(1): 167-174.

Oram, N. and L. Cuddy (1995). "Responsiveness of Western adults to pitch-distributional information in melodic sequences." Psychological Research **57**: 103-118.

Orman, E. K. (2011). "The effect of listening to specific musical genre selections on measures of heart rate variability." Update: Applications of Research in Music Education **30**(1): 64-69.

Osmanski, M. S., X. Song, et al. (2016). "Frequency discrimination in the common marmoset (Callithrix jacchus)." Hearing research **341**: 1-8.

Osmanski, M. S., X. Song, et al. (2013). "The role of harmonic resolvability in pitch perception in a vocal nonhuman primate, the common marmoset (Callithrix jacchus)." Journal of Neuroscience **33**(21): 9161-9168.

Ouattara, K., A. Lemasson, et al. (2009). "Campbell's monkeys concatenate vocalizations into context-specific call sequences." Proceedings of the National Academy of Sciences **106**(51): 22026-22031.

Oullier, O., G. C. De Guzman, et al. (2008). "Social coordination dynamics: Measuring human bonding." Social neuroscience **3**(2): 178-192.

Oullier, O., K. Jantzen, et al. (2004). "Neural substrates of real and imagined sensorimotor coordination." Cerebral cortex **15**(7): 975-985.

Ouyang, I. C. and E. Kaiser (2015). "Prosody and information structure in a tone language: an investigation of Mandarin Chinese." Language, Cognition and Neuroscience **30**(1-2): 57-72.

Owings, D. and D. Hennesey (1984). The importance of variation in sciurid visual and vocal communication. Biology of ground-dwelling squirrels: Annual cycles, behavioural ecology and sociality. J. Murie and G. Michener. Lincoln, University of Nebraska Press**:** 202-247.

Owings, D. and E. Morten (1998). Animal vocal communication: A new approach. Cambridge, Cambridge University Press.

Owren, M. and R. Bernacki (1988). "The acoustic features of alarm calls in vervet monkeys (Cercopithecus aethiops)." Journal of the Acoustical Society of America **5**: 1927-1935.

Owren, M. J., R. T. Amoss, et al. (2011). "Two organizing principles of vocal production: Implications for nonhuman and human primates." American Journal of Primatology **73**(6): 530-544.

Oxenham, A. J. (2012). "Pitch perception." Journal of Neuroscience **32**(39): 13335-13338.

Oxenham, A. J., C. Micheyl, et al. (2011). "Pitch perception beyond the traditional existence region of pitch." Proceedings of the National Academy of Sciences **108**(18): 7629-7634.

Özdemir, E., A. Norton, et al. (2006). "Shared and distinct neural correlates of singing and speaking." Neuroimage **33**(2): 628-635.

Paabo, S. (1995). "The Y chromosome and the origin of all of us (men)." Science **268**: 1141.

Pabon, P., S. Ternstrom, et al. (2011). "Fourier descriptor analysis and unification of voice range profile contours: method and applications." Journal of Speech, Language, and Hearing Research **54**(3): 755-776.

Page, S. C., S. H. Hulse, et al. (1989). "Relative pitch perception in the European starling (Sturnus vulgaris): further evidence for an elusive phenomenon." Journal of Experimental Psychology: Animal Behavior Processes **15**(2): 137.

Palastanga, N., D. Field, et al. (2006). Anatomy and human movement: structure and function, Elsevier Health Sciences.

Pally, R. and D. D. Olds (2000). The mind-brain relationship, Other Press, LLC.

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

Panksepp, J. (2009). "The emotional antecedents to the evolution of music and language." Musicae Scientiae **13**(2_suppl): 229-259.

Panksepp, J. (2010). "Affective consciousness in animals: perspectives on dimensional and primary process emotion approaches." Proceedings of the Royal Society of London B: Biological Sciences **277**(1696): 2905-2907.

Panksepp, J. (2011). "The basic emotional circuits of mammalian brains: do animals have affective lives?" Neuroscience & Biobehavioral Reviews **35**(9): 1791-1804.

Panksepp, J. and C. Trevarthen (2009). The neuroscience of emotion in music. Communicative musicality: Exploring the basis of human companionship. S. M. C. Trevarthen. New York, Oxford University Press.**:** 105-146.

Panzarella, R. (1980). "The phenomenology of aesthetic peak experiences." Journal of Humanistic Psychology **20**: 69-85.

Papousek, H. (1996). Musicality in infant research. Musical beginnings. I. Deliege and J. Sloboda (eds.). Oxford, Oxford University Press.

Papousek, H., M. Papousek, et al. (1985). The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. Social perception in infants. T. Field and N. Fox. Norwood, NJ, Ablex**:** 269-297.

Papousek, M. (1996). Intuitive Parenting. Musical Beginnings. I. Deliege and J. Sloboda. Oxford, Oxford University Press**:** 88-111.

Papousek, M. and H. Papousek (1991). Innate and cultural guidance of infants' integrative competencies. Cultural approaches to parenting. M. Bornstein. Hillsdale, NJ, Erlbaum.

Papousek, M., H. Papousek, et al. (1991). "The meanings of melodies in motherese in tone and stress languages." Infant Behaviour and Development **14**: 415-440.

Parlitz, D. and M. Pangert (1998). Short and medium motor responses to pitch shift: Latency measurements of the professional musician's audio-motor loop for intonation. 16th international congress on acoustics and the 137th meeting of the acoustical society of America, Seattle, WA.

Pastore, R. E. and S. M. Farrington (1996). "Measuring the difference limen for identification of order of onset for complex auditory stimuli." Attention, Perception, & Psychophysics **58**(4): 510-526.

Patel, A. and S. Demorest (2013). "Comparative music cognition: cross-species and cross-cultural studies." The psychology of music **3**: 647-681.

Patel, A. and J. Iversen (2006). A non-human animal can drum a steady beat on a musical instrument. Proceedings of the 9th International Conference on Music Perception and Cognition.

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

Patel, A. D. (2010). "Music, biological evolution, and the brain." Emerging disciplines: 91-144.

Patel, A. D. (2010). Music, language, and the brain, Oxford University Press.

Patel, A. D. (2017). "Why Doesn't a Songbird (the European Starling) Use Pitch to Recognize Tone Sequences? The Informational Independence Hypothesis." Comparative Cognition & Behavior Reviews **12**.

Patel, A. D. and E. Balaban (2001). "Human pitch perception is reflected in the timing of stimulus-related cortical activity." Nature neuroscience **4**(8): 839.

Patel, A. D. and J. R. Iversen (2014). "The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis." Frontiers in systems neuroscience **8**.

Patel, A. D., J. R. Iversen, et al. (2009). "Experimental evidence for synchronization to a musical beat in a nonhuman animal." Current Biology **19**(10): 827-830.

Patel, A. D., J. R. Iversen, et al. (2005). "The influence of metricality and modality on synchronization with a beat." Experimental brain research **163**(2): 226-238.

Patel, A. D., I. Peretz, et al. (1998). "Processing prosodic and musical patterns: a neuropsychological investigation." Brain and language **61**(1): 123-144.

Patel, S., C. Nishimura, et al. (2014). "Understanding the mechanisms underlying voluntary responses to pitch-shifted auditory feedback." The Journal of the Acoustical Society of America **135**(5): 3036-3044.

Patterson, R. D., S. Uppenkamp, et al. (2002). "The processing of temporal pitch and melody information in auditory cortex." Neuron **36**(4): 767-776.

Paus, T., M. Petrides, et al. (1993). "Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study." Journal of neurophysiology **70**(2): 453-469.

Pearce, E. (2014). "Modelling mechanisms of social network maintenance in hunter–gatherers." Journal of archaeological science **50**: 403-413.

Pearce, E., J. Launay, et al. (2015). "The ice-breaker effect: singing mediates fast social bonding." Royal Society Open Science **2**(10): 150221.

Pearce, E., J. Launay, et al. (2017). "Tuning in to others: exploring relational and collective bonding in singing and non-singing groups over time." Psychology of Music **45**(4): 496-512.

Penagos, H., J. R. Melcher, et al. (2004). "A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging." Journal of Neuroscience **24**(30): 6810-6815.

Peng, Y.-C., T.-W. Lu, et al. (2011). "Immediate effects of therapeutic music on loaded sit-to-stand movement in children with spastic diplegia." Gait & Posture **33**(2): 274-278.

Perani, D., M. C. Saccuman, et al. (2010). "Functional specializations for music processing in the human newborn brain." Proceedings of the National Academy of Sciences **107**(10): 4758-4763.

Peretz, I. (1996). "Can we lose memories for music? The case of music agnosia in a nonmusician." Journal of Cognitive Neurosciences **8**(6): 481-496.

Peretz, I. (2001). Listen to the brain: A biological perspective on musical emotions. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press.

Peretz, I. (2001a). The biological foundations of music. Language brain cognitive development: Essays in honor of Jacques Mehler. E. Dupoux. Cambridge, The Bradford Book.

Peretz, I. (2001b). "Brain specialization for music: New evidence from congenital amusia." Annals of the New York Academy of Sciences **930**: 153-165.

Peretz, I. (2003). Brain specialization for music: new evidence from congenital amusia. The cognitive neuroscience of music. I. Peretz and R. Zatorre. Oxford, Oxford University Press.

Peretz, I., Ed. (2006). Journal of Cognition, special issue: The nature of music. **100**(1).

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

Peretz, I. (2016). "Neurobiology of Congenital Amusia." Trends in cognitive sciences **20**(11): 857-867.

Peretz, I., J. Ayotte, et al. (2002). "Congenital amusia: a disorder of fine-grained pitch discrimination." Neuron **33**(2): 185-191.

Peretz, I., S. Belleville, et al. (1997). "Dissociation entre musique et langage apres atteinte cerebrale: un nouveau cas d'amusie sans aphasie." Canadian Journal of Experimental Psychology **51**(4): 354-368.

Peretz, I., E. Brattico, et al. (2009). "The amusic brain: in tune, out of key, and unaware." Brain **132**(5): 1277-1286.

Peretz, I., E. Brattico, et al. (2005). "Abnormal electrical brain responses to pitch in congenital amusia." Annals of neurology **58**(3): 478-482.

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

Peretz, I., S. Cummings, et al. (2007). "The genetics of congenital amusia (tone deafness): a family-aggregation study." The American Journal of Human Genetics **81**(3): 582-588.

Peretz, I., L. Gagnon, et al. (1998). "Music and emotion: perceptual determinants, immediacy, and isolation after brain damage." Cognition **68**(2): 111-141.

Peretz, I., L. Gagnon, et al. (2004). "Singing in the brain: Insights from cognitive neuropsychology." Music Perception **21**(3): 111-141.

Peretz, I. and K. L. Hyde (2003). "What is specific to music processing? Insights from congenital amusia." Trends in cognitive sciences **7**(8): 362-367.

Peretz, I., R. Kolinsky, et al. (1994). "Functional dissociations following bilateral lesions of auditory cortex." Brain **117**: 1283-1301.

Peretz, I. and R. Zatorre, Eds. (2003). The cognitive neuroscience of music. Oxford, Oxford University Press.

Peretz, I. and R. Zatorre (2005). "Brain organization for music processing." Annual Review of Psychology **56**: 89-114.

Perlman, M. and N. Clark (2015). "Learned vocal and breathing behavior in an enculturated gorilla." Animal cognition **18**(5): 1165-1179.

Perry, D. W., R. J. Zatorre, et al. (1999). "Localization of cerebral activity during simple singing." Neuroreport **10**(18): 3979-3984.

Persson, R. S. (2000). "Survival of the fittest or the most talented?: deconstructing the myth of the musical maestro." Journal of Secondary Gifted Education **12**(1): 25-38.

Peter, B., C. Stoel-Gammon, et al. (2008). “Octave equivalence as an aspect of stimulus response similarity during nonword and sentence imitations in young children”. Speech Prosody 731-734.

Peterson, S., P. Fox, et al. (1988). "Positron emission tomographic studies of the cortical anatomy of single-word processing." Nature **331**: 585-589.

Petrides, M., G. Cadoret, et al. (2005). "Orofacial somatomotor responses in the macaque monkey homologue of Broca's area." Nature **435**(7046): 1235-1238.

Petrides, M. and D. N. Pandya (2009). "Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey." PLoS biology **7**(8): e1000170.

Pezzulo, G., M. V. Butz, et al. (2008). From sensorimotor to higher-level cognitive processes: An introduction to anticipatory behavior systems. Workshop on Anticipatory Behavior in Adaptive Learning Systems, Springer.

Pezzulo, G. and C. Castelfranchi (2009). "Thinking as the control of imagination: a conceptual framework for goal-directed systems." Psychological Research PRPF **73**(4): 559-577.

Pfeifer, J. and S. Hamann (2015). "Revising the diagnosis of congenital amusia with the Montreal Battery of Evaluation of Amusia." Frontiers in human neuroscience **9**.

Pfordresher, P. Q. and S. Brown (2007). "Poor-pitch singing in the absence of" tone deafness"." Music Perception: An Interdisciplinary Journal **25**(2): 95-115.

Pfordresher, P. Q. and S. Brown (2017). "Vocal mistuning reveals the origin of musical scales." Journal of Cognitive Psychology **29**(1): 35-52.

Philips-Silver, J. and L. Trainor (2005). "Feeling the beat: Movement influences infant rhythm perception." Science **308**: 1430.

Phillips-Silver, J., P. Toiviainen, et al. (2013). "Amusic does not mean unmusical: Beat perception and synchronization ability despite pitch deafness." Cognitive neuropsychology **30**(5): 311-331.

Phillips-Silver, J., P. Toiviainen, et al. (2011). "Born to dance but beat deaf: a new form of congenital amusia." Neuropsychologia **49**(5): 961-969.

Phillips-Silver, J. and L. J. Trainor (2007). "Hearing what the body feels: Auditory encoding of rhythmic movement." Cognition **105**(3): 533-546.

Phillips-Silver, J. and L. J. Trainor (2008). "Vestibular influence on auditory metrical interpretation." Brain and cognition **67**(1): 94-102.

Piccirilli, M., T. Sciarma, et al. (2000). "Modularity of music: Evidence from a case of pure amusia." Journal of Neurology, Neurosurgery Psychiatry **69**: 541-545.

Pickens, J. and L. Bahrick (1995). "Infants' discrimination of bimodal events on the basis of rhythm and tempo." British Journal of Developmental Psychology **13**: 223-236.

Pickens, J. and L. Bahrick (1997). "Do infants perceive invariant tempo and rhythm in auditory-visual events?" Infant Behaviour and Development **20**: 349-357.

Pignatiello, M., C. Camp, et al. (1989). "A psychophysiological comparison of the Velten and musical mood induction techniques." Journal of Music Therapy **26**: 140-154.

Pike, A. (1972). "A phenomenological analysis of emotional experience in music." Journal of Research in Music Education **20**: 262-267.

Pike, K. L. (1945). "The intonation of American English." University of Michigan Press.

Pinker, S. (1994). The language instinct. London, Allen Lane.

Pinker, S. (1997). How the mind works. New York, W.W. Norton.

Pisanski, K., V. Cartei, et al. (2016). "Voice modulation: A window into the origins of human vocal control?" Trends in cognitive sciences **20**(4): 304-318.

Plack, C. J., D. Barker, et al. (2014). "Pitch coding and pitch processing in the human brain." Hearing research **307**: 53-64.

Plack, C. J. and A. J. Oxenham (2005). The psychophysics of pitch. Pitch, Springer**:** 7-55.

Plack, C. J., A. J. Oxenham, et al. (2006). Pitch: neural coding and perception, Springer Science & Business Media.

Plantinga, J. and L. Trainor (2005). "Memory for melody: Infants use a relative pitch code." Cognition **98**: 1-11.

Plantinga, J. and S. E. Trehub (2014). "Revisiting the innate preference for consonance." Journal of experimental psychology: Human perception and performance **40**(1): 40.

Ploeger, A. (2017). "Nomological Networks of Evidence." Encyclopedia of Evolutionary Psychological Science: 1-4.

Plomp, R. and W. J. M. Levelt (1965). "Tonal consonance and critical bandwidth." The Journal of the Acoustical Society of America **38**(4): 548-560.

Plutchik, R. (1994). The psychology and biology of emotion. New York, Harper Collins.

Podlipniak, P. (2016). "The evolutionary origin of pitch centre recognition." Psychology of Music **44**(3): 527-543.

Poeppel, D. (2003). "The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’." Speech communication **41**(1): 245-255.

Pöppel, E. (1973). "Influence of pause duration on the reproduction of a 2-second interval." Bulletin of the Psychonomic Society **2**(5): 291-292.

Pöppel, E. (1978). Time perception. Perception, Springer**:** 713-729.

Pöppel, E. (1996). "Reconstruction of subjective time on the basis of hierarchically organized processing system." Advances in psychology **115**: 165-185.

Pöppel, E. (1997). "A hierarchical model of temporal perception." Trends in cognitive sciences **1**(2): 56-61.

Pöppel, E. and M. Wittmann (1999). "Time in the mind." The MIT encyclopedia of the cognitive sciences: 841-843.

Porges, S. W. (2001). "The polyvagal theory: phylogenetic substrates of a social nervous system." International Journal of Psychophysiology **42**(2): 123-146.

Porges, S. W. (2003). "The polyvagal theory: Phylogenetic contributions to social behavior." Physiology & behavior **79**(3): 503-513.

Porges, S. W. (2007). "The polyvagal perspective." Biological psychology **74**(2): 116-143.

Porges, S. W. (2009). "The polyvagal theory: new insights into adaptive reactions of the autonomic nervous system." Cleveland Clinic journal of medicine **76**(Suppl 2): S86.

Porges, S. W. (2011). The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation (Norton Series on Interpersonal Neurobiology), WW Norton & Company.

Post, J., Ed. (2006). Ethnomusicology: a contemporary reader. New York, Routledge.

Post, J. C. (2013). Ethnomusicology: a contemporary reader, Routledge.

Povinelli, D. J. and T. J. Eddy (1996). "Chimpanzees: Joint visual attention." Psychological science **7**(3): 129-135.

Povinelli, D. J. and D. K. O'Neill (2000). Do chimpanzees use their gestures to instruct each other? Understanding other minds: Perspectives from developmental cognitive neuroscience. H. T.-

F. S. Baron-Cohen, & D. J. Cohen. New York, Oxford University Press**:** 459-487.

Power, M. and T. Dalgleish (1997). Cognition and emotion. From order to disorder. Hove, UK, Psychology Press.

Praamstra, P., D. Kourtis, et al. (2006). "Neurophysiology of implicit timing in serial choice reaction-time performance." Journal of Neuroscience **26**(20): 5448-5455.

Praamstra, P., M. Turgeon, et al. (2003). "Neurophysiological correlates of error correction in sensorimotor-synchronization." Neuroimage **20**(2): 1283-1297.

Pras, A., M. F. Schober, et al. (2017). "What About Their Performance Do Free Jazz Improvisers Agree Upon? A Case Study." Frontiers in psychology **8**: 966.

Prassas, S., M. Thaut, et al. (1997). "Effect of auditory rhythmic cuing on gait kinematic parameters of stroke patients." Gait & Posture **6**(3): 218-223.

Pressing, J. (1998). "Error correction processes in temporal pattern production." Journal of Mathematical Psychology **42**(1): 63-101.

Pressing, J. (1999). "The referential dynamics of cognition and action." Psychological review **106**(4): 714.

Prinz, W. (1990). A common coding approach to perception and action. Relationships between perception and action, Springer**:** 167-201.

Prosen, C., D. Moody, et al. (1990). "Frequency discrimination in the monkey." The Journal of the Acoustical Society of America **88**(5): 2152-2158.

Puschmann, S., S. Uppenkamp, et al. (2010). "Dichotic pitch activates pitch processing centre in Heschl's gyrus." Neuroimage **49**(2): 1641-1649.

Rabinowitch, T.-C. (2015). "How, Rather Than What Type of, Music Increases Empathy." Empirical Musicology Review **10**(1-2): 96-98.

Rabinowitch, T.-C. (2017). "2 Synchronisation–a musical substrate for positive social interaction and empathy." Music and Empathy: 89.

Rabinowitch, T.-C., I. Cross, et al. (2013). "Long-term musical group interaction has a positive influence on empathy in children." Psychology of Music **41**(4): 484-498.

Rabinowitch, T.-C. and A. Knafo-Noam (2015). "Synchronous rhythmic interaction enhances children's perceived similarity and closeness towards each other." PloS one **10**(4): e0120878.

Racette, A. and I. Peretz (in press). "Learning lyrics: To sing or not to sing?" Memory & Cognition.

Racine, E., O. Bar-Ilan, et al. (2005). "fMRI in the public eye." Nature Reviews Neuroscience **6**(2): 159-164.

Rakoczy, H., M. Tomasello, et al. (2005). "How children turn objects into symbols: A cultural learning account." Symbol use and symbol representation. New York: Erlbaum: 67-97.

Ramachandran, V. S. (2012). The tell-tale brain: A neuroscientist's quest for what makes us human, WW Norton & Company.

Ramus, F. (2002). "Acoustic correlates of linguistic rhythm: Perspectives." Proceedings of Speech Prosody 2002, Aix-en-Provence (pp. 115–120). Aix-en-Provence: Laboratoire Parole et Langage.

Ramus, F., E. Dupoux, et al. (2003). "The psychological reality of rhythm classes: Perceptual studies." In Proceedings of the 15th international congress of phonetic sciences (pp. 337–342). Barcelona.

Ramus, F., M. Nespor, et al. (1999). "Correlates of linguistic rhythm in the speech signal." Cognition **73**(3): 265-292.

Rand, A. and M. Ryan (1981). "The adaptive significance of a complex vocal repertoire in a neotropical frog." Zeitschrift für Tierpsychologie **57**: 209-214.

Randall, J. A. (1997). "Species-specific footdrumming in kangaroo rats: Dipodomys ingens, D. deserti, D. spectabilis." Animal Behaviour **54**(5): 1167-1175.

Rao, S. M., D. L. Harrington, et al. (1997). "Distributed neural systems underlying the timing of movements." Journal of Neuroscience **17**(14): 5528-5535.

Rappaport, R. A. (1999). Ritual and Religion in the Making of Humanity, Cambridge University Press.

Ratner, N. and J. Bruner (1978). "Games, social exchange and the acquisition of language." Journal of child language **5**(3): 391-401.

Rauschecker, J. P. and S. K. Scott (2009). "Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing." Nature neuroscience **12**(6): 718-724.

Ravignani, A., W. T. Fitch, et al. (2016). "What pinnipeds have to say about human speech, music, and the evolution of rhythm." Frontiers in neuroscience **10**.

Raymond, J. E., K. L. Shapiro, et al. (1992). "Temporary suppression of visual processing in an RSVP task: An attentional blink?" Journal of experimental psychology: Human perception and performance **18**(3): 849.

Read, D. W. (2008). "Working memory: a cognitive limit to non-human primate recursive thinking prior to hominid evolution." Evolutionary Psychology **6**(4): 147470490800600413.

Rebuschat, P., M. Rohrmeier, et al. (2011). Language and Music as Cognitive Systems, OUP Oxford.

Recanzone, G. H., W. M. Jenkins, et al. (1991). "A behavioral frequency discrimination paradigm for use in adult primates." Behavior Research Methods, Instruments, & Computers **23**(3): 357-369.

Reddish, P., J. Bulbulia, et al. (2014). "Does synchrony promote generalized prosociality?" Religion, Brain & Behavior **4**(1): 3-19.

Reddish, P., R. Fischer, et al. (2013). "Let's dance together: synchrony, shared intentionality and cooperation." PloS one **8**(8): e71182.

Reddish, P., E. M. Tong, et al. (2016). "Collective synchrony increases prosociality towards non-performers and outgroup members." British Journal of Social Psychology **55**(4): 722-738.

Reich, C. M., J. S. Berman, et al. (2014). "Vocal synchrony in psychotherapy." Journal of Social and Clinical Psychology **33**(5): 481-494.

Reich, D. and D. Goldstein (1998). "Genetic evidence for a Paleolithic human population expansion in Africa." Proceedings of the National Academy of Sciences **95**(18): 8119-8123.

Reichard, U. H., C. Barelli, et al. (2016). The evolution of gibbons and siamang. Evolution of Gibbons and Siamang, Springer**:** 3-41.

Reichmuth, C. and C. Casey (2014). "Vocal learning in seals, sea lions, and walruses." Current opinion in neurobiology **28**: 66-71.

Remedios, R., N. Logothetis, et al. (2010). Vocal and non-vocal acoustic communication systems in the macaque brain. 40th Annual Meeting of the Society for Neuroscience (Neuroscience 2010).

Remedios, R., N. K. Logothetis, et al. (2009). "Monkey drumming reveals common networks for perceiving vocal and nonvocal communication sounds." Proceedings of the National Academy of Sciences **106**(42): 18010-18015.

Rendall, D., D. Cheney, et al. (2000). "Proximate factors mediating 'contact' calls in adult female baboons and their infants." Journal of Comparative Psychology **114**: 36-46.

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

Repp, B. (2004). "Comments on "Rapid motor adaptations to subliminal frequency shifts during syncopated rhythmic sensorimotor synchronisation" by M. Thaut & G. Kenyon." Human Movement Science **21**(3): 61-78.

Repp, B. H. (1993). "Music as motion: A synopsis of Alexander Truslit's (1938) Gestaltung und Bewegung in der Musik." Psychology of Music **21**(1): 48-72.

Repp, B. H. (1999). "Control of expressive and metronomic timing in pianists." Journal of Motor behavior **31**(2): 145-164.

Repp, B. H. (2000). "Compensation for subliminal timing perturbations in perceptual-motor synchronization." Psychological Research **63**(2): 106-128.

Repp, B. H. (2001). "Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization." Journal of experimental psychology: Human perception and performance **27**(3): 600.

Repp, B. H. (2001). "Processes underlying adaptation to tempo changes in sensorimotor synchronization." Human Movement Science **20**(3): 277-312.

Repp, B. H. (2002). "Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization." Journal of Experimental Psychology-Human Perception and Performance **28**(2): 410-430.

Repp, B. H. (2003). "Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision." Journal of Motor behavior **35**(4): 355-370.

Repp, B. H. (2004). "Comments on “Rapid motor adaptations to subliminal frequency shifts during syncopated rhythmic sensorimotor synchronization” by Michael H. Thaut and Gary P. Kenyon (Human Movement Science 22 [2003] 321–338)." Human Movement Science **23**(1): 61-77.

Repp, B. H. (2005). "Sensorimotor synchronization: a review of the tapping literature." Psychonomic bulletin & review **12**(6): 969-992.

Repp, B. H. (2008). "Metrical subdivision results in subjective slowing of the beat." Music Perception: An Interdisciplinary Journal **26**(1): 19-39.

Repp, B. H. (2008). "Perfect phase correction in synchronization with slow auditory sequences." Journal of Motor behavior **40**(5): 363-367.

Repp, B. H. (2010). "Sensorimotor synchronization and perception of timing: effects of music training and task experience." Human Movement Science **29**(2): 200-213.

Repp, B. H. (2011). "Tapping in synchrony with a perturbed metronome: The phase correction response to small and large phase shifts as a function of tempo." Journal of Motor behavior **43**(3): 213-227.

Repp, B. H. and R. Doggett (2007). "Tapping to a very slow beat: a comparison of musicians and nonmusicians." Music Perception: An Interdisciplinary Journal **24**(4): 367-376.

Repp, B. H. and P. E. Keller (2004). "Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness." Quarterly Journal of Experimental Psychology Section A **57**(3): 499-521.

Repp, B. H. and P. E. Keller (2008). "Sensorimotor synchronization with adaptively timed sequences." Human Movement Science **27**(3): 423-456.

Repp, B. H., P. E. Keller, et al. (2012). "Quantifying phase correction in sensorimotor synchronization: empirical comparison of three paradigms." Acta psychologica **139**(2): 281-290.

Repp, B. H., J. London, et al. (2005). "Production and synchronization of uneven rhythms at fast tempi." Music Perception: An Interdisciplinary Journal **23**(1): 61-78.

Repp, B. H., J. London, et al. (2008). "Phase correction in sensorimotor synchronization with nonisochronous sequences." Music Perception: An Interdisciplinary Journal **26**(2): 171-175.

Repp, B. H., J. London, et al. (2011). "Perception–production relationships and phase correction in synchronization with two-interval rhythms." Psychological Research **75**(3): 227-242.

Repp, B. H., J. London, et al. (2013). "Systematic distortions in musicians’ reproduction of cyclic three-interval rhythms." Music Perception: An Interdisciplinary Journal **30**(3): 291-305.

Repp, B. H. and G. P. Moseley (2012). "Anticipatory phase correction in sensorimotor synchronization." Human Movement Science **31**(5): 1118-1136.

Repp, B. H. and A. Penel (2002). "Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences." Journal of Experimental Psychology-Human Perception and Performance **28**(5): 1085-1099.

Repp, B. H. and Y.-H. Su (2013). "Sensorimotor synchronization: a review of recent research (2006–2012)." Psychonomic bulletin & review **20**(3): 403-452.

Reznikoff, I. and M. Dauvois (1988). "La dimension sonore des grottes ornees." Bulletin de la Societe Prehistorique Francaise **85**: 238-246.

Richardson, M. J., K. L. Marsh, et al. (2007). "Rocking together: Dynamics of intentional and unintentional interpersonal coordination." Human Movement Science **26**(6): 867-891.

Richter, D., J. Waiblinger, et al. (2000). "Thermoluminescence, electron spin resonance and C-14-dating of the late middle and early upper Paleolithic site of Geissenkloesterle cave in southern Germany." Journal of Archaeological Science **27**: 71-89.

Rickard, N. S., P. Appelman, et al. (2013). "Orchestrating life skills: The effect of increased school-based music classes on children's social competence and self-esteem." International Journal of Music Education **31**(3): 292-309.

Rickard, N. S., C. J. Bambrick, et al. (2012). "Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10–13-year-old students." International Journal of Music Education **30**(1): 57-78.

Ridderinkhof, K. R., W. P. Van Den Wildenberg, et al. (2004). "Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning." Brain and cognition **56**(2): 129-140.

Ridgway, S., D. Carder, et al. (2012). "Spontaneous human speech mimicry by a cetacean." Current Biology **22**(20): R860-R861.

Riecker, A., H. Ackermann, et al. (2000). "Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum." Neuroreport **11**(9): 1997-2000.

Riecker, A., K. Mathiak, et al. (2005). "fMRI reveals two distinct cerebral networks subserving speech motor control." Neurology **64**(4): 700-706.

Riest, C., A. B. Jorschick, et al. (2015). "Anticipation in turn-taking: mechanisms and information sources." Frontiers in psychology **6**.

Riganello, F., A. Candelieri, et al. (2010). "Heart rate variability, emotions, and music." Journal of Psychophysiology. Journal of Psychophysiology, **24**(2), 112-119.

Rigg, M. (1937). "Musical expression: An investigation of the theories of Erich Sorantin." Journal of Experimental Psychology **21**: 442-455.

Riggle, M. (2009). "A simpler explanation for vestibular influence on beat perception: no specialized unit needed." Empirical Musicological Review **4**: 14-16.

Rilling, J. K., M. F. Glasser, et al. (2008). "The evolution of the arcuate fasciculus revealed with comparative DTI." Nature neuroscience **11**(4): 426.

Roach, P. (1982). "On the distinction between 'stress-timed'and 'syllable-timed'languages." Linguistic controversies **73**: 79.

Roberts, S. G., F. Torreira, et al. (2015). "The effects of processing and sequence organization on the timing of turn taking: a corpus study." Frontiers in psychology **6**.

Robledo, J. P., S. Hawkins, et al. (2016). Pitch-interval analysis of 'periodic'and 'aperiodic'Question+ Answer pairs. Speech Prosody 2016, York.

Robledo, J. P., E. Hurtado, et al. (2016). "Music intervals in speech: Psychological disposition modulates ratio precision among interlocutors' nonlocal f0 production in real-time dyadic conversation." Psychology of Music **44**(6): 1404-1418.

Rock, A., L. Trainor, et al. (1999). "Distinctive messages in infant-directed lullabies and play songs." Developmental Psychology **35**(2): 527-534.

Rodriguez, C. and P. Palacios (2007). "Do private gestures have a self-regulatory function?: a case study." Infant behavior and development **30**(2): 180-194.

Roe, K. (1985). "Swedish youth and music: Listening patterns and motivations." Communication Research **12**: 353-362.

Roederer, J. (1984). "The search for the survival function of music." Music Perception **1**: 350-356.

Roerdink, M., P. J. Bank, et al. (2011). "Walking to the beat of different drums: Practical implications for the use of acoustic rhythms in gait rehabilitation." Gait & Posture **33**(4): 690-694.

Roerdink, M., C. J. Lamoth, et al. (2007). "Gait coordination after stroke: benefits of acoustically paced treadmill walking." Physical Therapy **87**(8): 1009-1022.

Rogers Montgomery, C. and M. G. Clarkson (1997). "Infants' pitch perception: masking by low-and high-frequency noises." The Journal of the Acoustical Society of America **102**(6): 3665-3672.

Rohrmeier, M. and P. Rebuschat (2012). "Implicit learning and acquisition of music." Topics in cognitive science **4**(4): 525-553.

Rohrmeier, M. and R. Widdess (2017). "Incidental Learning of Melodic Structure of North Indian Music." Cognitive science **41**(5): 1299-1327.

Rohrmeier, M., W. Zuidema, et al. (2015). "Principles of structure building in music, language and animal song." Phil. Trans. R. Soc. B **370**(1664): 20140097.

Roseman, M. (1991). Healing sounds from the Malaysian rainforest: Temiar music and medicine, Univ of California Press.

Ross, B. and S. Knight (2017). "Reports of equitonic scale systems in African musical traditions and their implications for cognitive models of pitch organization." Musicae Scientiae: 1029864917736105.

Ross, C. and M. Ravosa (1993). "Basicranial flexion, relative brain size, and facial kyphosis in nonhuman primates." American Journal of Physical Anthropology **91**: 305-314.

Ross, H. S. and S. P. Lollis (1987). "Communication within infant social games." Developmental Psychology **23**(2): 241.

Ross, J. M., J. R. Iversen, et al. (2016). "Motor simulation theories of musical beat perception." Neurocase **22**(6): 558-565.

Rouget, G. (1985). Music and trance. Chicago, University of Chicago Press.

Rouget, G. (2004). "Music with Effects: Musicking to Survive." L'Homme(3): 27-52.

Rouse, A. A., P. F. Cook, et al. (2016). "Beat keeping in a sea lion as coupled oscillation: implications for comparative understanding of human rhythm." Frontiers in neuroscience **10**.

Ruggero, M. A. (1992). "Responses to sound of the basilar membrane of the mammalian cochlea." Current opinion in neurobiology **2**(4): 449-456.

Runciman, W. (1998). "The selectionist paradigm and its implications for sociology." Sociology **32**: 163-188.

Russell, J. (1980). "A circumplex model of affect." Journal of Personality and Social Psychology **39**: 1161-1178.

Russell, J. A. (2003). "Core affect and the psychological construction of emotion." Psychological review **110**(1): 145.

Ryan, M. (1980). "Female mate choice in a neotropical frog." Science **209**: 523-525.

Ryan, M. (1985). The tungara frog: A study in sexual selection and communication. Chicago, Chicago University Press.

Ryan, M. and E. Brenowitz (1985). "The role of body size, phylogeny, and ambient noise in the evolution of bird song." The American Journal of Nature **126**: 87-100.

Ryan, M., S. Phelps, et al. (2001). "How evolutionary history shapes recognition mechanisms." Trends in Cognitive Science **5**: 143-148.

Ryan, R. and J. La Guardia (2000). "What is being optimized over development? A self-determination theory perspective on basic psychological needs across the life span." In S. Qualls & R. Abeles (Eds.), Dialogues on psychology and aging. Washington, DC: American Psychological Association.

Ryan, R. M. and E. L. Deci (2000). "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." American psychologist **55**(1): 68.

Ryan, R. M. and E. L. Deci (2001). "On happiness and human potentials: A review of research on hedonic and eudaimonic well-being." Annual review of psychology **52**(1): 141-166.

Saarikallio, S. (2016). "Music as emotion regulation." Musicae Scientiae **20**(1): 10-10.

Saarikallio, S. and J. Erkkilä (2007). "The role of music in adolescents' mood regulation." Psychology of Music **35**(1): 88-109.

Sachs, C. (1943). The rise of music in the ancient world: East and West. New York, Norton.

Sacks, H., E. A. Schegloff, et al. (1974). "A simplest systematics for the organization of turn-taking for conversation." language: 696-735.

Saffron, J. (2001). "Words in a sea of sounds: The output of infant statistical learning." Cognition **81**: 149-169.

Saffron, J. (2003). "Absolute pitch in infancy and adulthood: The role of tonal structure." Developmental Science **6**: 35-45.

Saffron, J. and G. Griegentrog (2001). "Absolute pitch in infant auditory learning of tone sequences by human infants and adults." Cognition **70**: 27-52.

Saffron, J., K. Reeck, et al. (2005). "Changing the tune: The structure of the input affects infants' use of absolute and relative pitch." Developmental Science **8**: 1-7.

Safranek, M. G., G. F. Koshland, et al. (1982). "Effect of auditory rhythm on muscle activity." Physical Therapy **62**(2): 161-168.

Saito, Y., K. Ishii, et al. (2006). "Cerebral networks for spontaneous and synchronized singing and speaking." Neuroreport **17**(18): 1893-1897.

Saleh, M., J. Reimer, et al. (2010). "Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues." Neuron **65**(4): 461-471.

Santos, L. R. and M. D. Hauser (1999). "How monkeys see the eyes: Cotton-top tamarins' reaction to changes in visual attention and action." Animal cognition **2**(3): 131-139.

Sanvito, S., F. Galimberti, et al. (2007). "Observational evidences of vocal learning in southern elephant seals: a longitudinal study." Ethology **113**(2): 137-146.

Sarter, M., B. Givens, et al. (2001). "The cognitive neuroscience of sustained attention: where top-down meets bottom-up." Brain research reviews **35**(2): 146-160.

Satoh, M., N. Kato, et al. (2016). "A case of musical anhedonia due to right putaminal hemorrhage: a disconnection syndrome between the auditory cortex and insula." Neurocase **22**(6): 518-525.

Satoh, M., T. Nakase, et al. (2011). "Musical anhedonia: selective loss of emotional experience in listening to music." Neurocase **17**(5): 410-417.

Satoh, M., K. Takeda, et al. (2005). "A case of amusia caused by the infarction of anterior portion of bilateral temporal lobes." Cortex **41**: 77-83.

Saville, B., Y. Kohli, et al. (1998). "mtDNA recombination in a natural population." Proceedings of the National Academy of Sciences **95**: 1331-1335.

Schachner, A., T. F. Brady, et al. (2009). "Spontaneous motor entrainment to music in multiple vocal mimicking species." Current Biology **19**(10): 831-836.

Schachner, A. and L. Garvin (2010). Does synchrony really affect social variables? Effects on cooperation, conformity may not be robust. Poster presented at the International Conference on Music Perception and Cognition, Seattle, WA.

Schaller, G. E. (1963). "The mountain gorilla: ecology and behavior." University of Chicago Press.

Schellenberg, E. and L. Trainor (1996). "Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners." Journal of the Acoustical Society of America **100**: 3321-3328.

Schellenberg, E. and S. Trehub (1994). "Frequency ratios and the discrimination of melodic intervals." Perception and Psychophysics **56**: 472-478.

Schellenberg, E. and S. Trehub (1996a). "Children's discrimination of melodic intervals." Developmental Psychology **32**: 1039-1050.

Schellenberg, E. and S. Trehub (1996b). "Natural musical intervals: Evidence from infant listeners." Psychological Science **7**: 272-277.

Schellenberg, E. and S. Trehub (1999). "Culture-general and culture-specific factors in the discrimination of melodies." Journal of Experimental Child Psychology **74**: 107-127.

Schellenberg, E. and S. Trehub (1999). "Good pitch memory is widespread." Psychological Science **14**: 262-266.

Schellenberg, E. G. (2004). "Music lessons enhance IQ." Psychological science **15**(8): 511-514.

Schellenberg, E. G. (2006). "Long-term positive associations between music lessons and IQ." Journal of Educational Psychology **98**(2): 457.

Schellenberg, E. G., K. A. Corrigall, et al. (2015). "Group music training and children's prosocial skills." PloS one **10**(10): e0141449.

Schellenberg, E. G. and M. Mankarious (2012). "Music training and emotion comprehension in childhood." Emotion **12**(5): 887.

Schellenberg, E. G. and S. E. Trehub (2003). "Good pitch memory is widespread." Psychological science **14**(3): 262-266.

Schellenberg, E. G. and M. W. Weiss (2013). Music and cognitive abilities. The Psychology of Music (Third Edition), Elsevier**:** 499-550.

Scherer, K. (1986). "Vocal affect expression: A review and a model for future research." Psychological Bulletin **99**(2): 143-165.

Scherer, K. (1993). "Neuroscience projections to current debates in emotion psychology." Cognition and Emotion **7**: 1-41.
Scherer, K. (1995). "Expressoin of emotion in voice and music." Journal of Voice **9**: 235-248.

Scherer, K. (1999). Appraisal theories. Handbook of cognition and emotion. T. Dagleish and M. Power. Chichester, UK, Wiley**:** 637-636.

Scherer, K., T. Johnstone, et al. (2003). Vocal expression of emotion. Handbook of affective sciences. R.Davidson, K. Scherer and H. Goldsmith. Oxford, Oxford University Press.

Scherer, K. and J. Oshinsky (1977). "Cue utilization in emotion attribution from auditory stimuli." Motivation and Emotion **1**: 331-346.

Scherer, K., A. Schorr, et al., Eds. (2001). Appraisal processes in emotion: Theory, method, and research. Oxford, Oxford University Press.

Scherer, K. R. (2000). "Psychological models of emotion." The neuropsychology of emotion **137**(3): 137-162.

Scherer, K. R. (2009). "The dynamic architecture of emotion: Evidence for the component process model." Cognition and emotion **23**(7): 1307-1351.

Scherer, K. R. and M. R. Zentner (2001). "Emotional effects of music: Production rules." Music and emotion: Theory and research: 361-392.

Schiavio, A. (2015). "Action, enaction, inter (en) action." Empirical Musicology Review **9**(3-4): 254-262.

Schiavio, A. and S. Høffding (2015). "Playing together without communicating? A pre-reflective and enactive account of joint musical performance." Musicae Scientiae **19**(4): 366-388.

Schiavio, A., D. Menin, et al. (2014). "Music in the flesh: Embodied simulation in musical understanding." Psychomusicology: Music, Mind, and Brain **24**(4): 340.

Schiavio, A., D. van der Schyff, et al. (2016). "Enacting musical emotions. sense-making, dynamic systems, and the embodied mind." Phenomenology and the cognitive sciences: 1-25.

Schimmack, U. and A. Grob (2000). "Dimensional models of core affect: A qualitative comparison by means of structural equation modeling." European Journal of Personality **14**: 325-245.

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

Schmidt, R. C. and B. O'Brien (1997). "Evaluating the dynamics of unintended interpersonal coordination." Ecological Psychology **9**(3): 189-206.

Schmidt, R. C. and M. J. Richardson (2008). Dynamics of interpersonal coordination. Coordination: Neural, behavioral and social dynamics, Springer**:** 281-308.

Schmitt, D. P. and J. J. Pilcher (2004). "Evaluating evidence of psychological adaptation: How do we know one when we see one?" Psychological science **15**(10): 643-649.

Schneider, P., V. Sluming, et al. (2005). "Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference." Nature neuroscience **8**(9): 1241-1247.

Schneider, S., P. Schönle, et al. (2007). "Using musical instruments to improve motor skill recovery following a stroke." Journal of neurology **254**(10): 1339-1346.

Schober, M. F. and N. Spiro (2014). "Jazz improvisers' shared understanding: a case study." Frontiers in psychology **5**: 808.

Schober, M. F. and N. Spiro (2016). "Listeners' and performers' shared understanding of jazz improvisations." Frontiers in psychology **7**: 1629.

Schon, D., B. Lorber, et al. (2004). "A selective deficit in the production of exact musical intervals following right-hemisphere damage." Cognitive Neuropsychology **21**(7): 773-784.

Schöner, G. (2002). "Timing, clocks, and dynamical systems." Brain and cognition **48**(1): 31-51.

Schöner, G. and J. Kelso (1988). "A synergetic theory of environmentally-specified and learned patterns of movement coordination." Biological cybernetics **58**(2): 71-80.

Schönwiesner, M. and R. J. Zatorre (2008). "Depth electrode recordings show double dissociation between pitch processing in lateral Heschl’s gyrus and sound onset processing in medial Heschl’s gyrus." Experimental brain research **187**(1): 97-105.

Schore, A. (1994). Affect regulation and the origin of the self: The neurobiology of emotional development. Hillsdale, NJ, Erlbaum.

Schore, A. N. (2003). Affect Dysregulation and Disorders of the Self. (Norton Series on Interpersonal Neurobiology). WW Norton & Company.

Schore, A. N. (2015). Affect regulation and the origin of the self: The neurobiology of emotional development, Routledge.

Schouten, J. F., R. Ritsma, et al. (1962). "Pitch of the residue." The Journal of the Acoustical Society of America **34**(9B): 1418-1424.

Schreuder, M., L. Van Eerten, et al. (2005). Speaking in minor and major keys.

Schroeder, C. E. and P. Lakatos (2009). "Low-frequency neuronal oscillations as instruments of sensory selection." Trends in neurosciences **32**(1): 9-18.

Schubert, E. (1999). Measurement and time series analysis of emotion in music. Sydney, Australia, University of South Wales. **PhD**.

Schubert, E. (2001). Continuous measurement of self-report emotional response to music. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press**:** 393-414.

Schubert, E. (2013). "Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives." Frontiers in psychology **4**.

Schubotz, R. I., A. D. Friederici, et al. (2000). "Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI." Neuroimage **11**(1): 1-12.

Schubotz, R. I. and D. Y. von Cramon (2001). "Interval and ordinal properties of sequences are associated with distinct premotor areas." Cerebral cortex **11**(3): 210-222.

Schulz, G., M. Varga, et al. (2005). "Functional neuroanatomy of human vocalization: an H215O PET study." Cerebral Cortex **15**(12): 1835-1847.

Schulze, H.-H. (1992). The error correction model for the tracking of a random metronome: Statistical properties and an empirical test. Time, action and cognition, Springer**:** 275-286.

Schulze, H.-H., A. Cordes, et al. (2005). "Keeping synchrony while tempo changes: Accelerando and ritardando." Music Perception: An Interdisciplinary Journal **22**(3): 461-477.

Schulze, H.-H. and D. Vorberg (2002). "Linear phase correction models for synchronization: Parameter identification and estimation of parameters." Brain and cognition **48**(1): 80-97.

Schusterman, R. J. (2008). "Vocal learning in mammals with special emphasis on pinnipeds." The evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication: 41-70.

Schusterman, R. J. and D. Kastak (2002). "Problem solving and memory." Marine mammal biology: An evolutionary approach: 371-387.

Schusterman, R. J. and C. Reichmuth (2008). "Novel sound production through contingency learning in the Pacific walrus (Odobenus rosmarus divergens)." Animal cognition **11**(2): 319-327.

Schwartze, M., P. E. Keller, et al. (2011). "The impact of basal ganglia lesions on sensorimotor synchronization, spontaneous motor tempo, and the detection of tempo changes." Behavioural brain research **216**(2): 685-691.

Schwartze, M., K. Rothermich, et al. (2012). "Functional dissociation of pre-SMA and SMA-proper in temporal processing." Neuroimage **60**(1): 290-298.

Schwartze, M., A. Tavano, et al. (2012). "Temporal aspects of prediction in audition: cortical and subcortical neural mechanisms." International Journal of Psychophysiology **83**(2): 200-207.

Scothern, P. (1992). The music-archaeology of the Paleolithic within its cultural setting. Department of Archaeology. Cambridge, University of Cambridge. **PhD**.

Scott, B. H., M. Mishkin, et al. (2012). "Monkeys have a limited form of short-term memory in audition." Proceedings of the National Academy of Sciences **109**(30): 12237-12241.

Searle, J. R. (1983). Intentionality: An essay in the philosophy of mind, Cambridge university press.

Searle, J. R. (1995). The construction of social reality, Simon and Schuster.

Sebanz, N., H. Bekkering, et al. (2006). "Joint action: bodies and minds moving together." Trends in cognitive sciences **10**(2): 70-76.

Seddon, F. A. and M. Biasutti (2009). "Modes of communication between members of a string quartet." Small group research **40**(2): 115-137.

Seeger, A. (1987). "Why Suya Sing." A musical anthropology of an Amazonian people.

Seeger, A. (2004). Why Suyá sing: a musical anthropology of an Amazonian people, University of Illinois Press.

Seesjärvi, E., T. Särkämö, et al. (2016). "The nature and nurture of melody: a twin study of musical pitch and rhythm perception." Behavior genetics **46**(4): 506-515.

Semjen, A., D. Vorberg, et al. (1998). "Getting synchronized with the metronome: Comparisons between phase and period correction." Psychological Research **61**(1): 44-55.

Sergeant, D. (1983). "The octave-percept or concept." Psychology of Music **11**(1): 3-18.

Sethares, W. A. (2005). Tuning, timbre, spectrum, scale, Springer Science & Business Media.

Seyfarth, R. and D. Cheney (2003). "Signalers and receivers in animal communication." Annual Review of Psychology **54**: 145-173.

Seyfarth, R. M., D. L. Cheney, et al. (1980). "Monkey responses to three different alarm calls: evidence of predator classification and semantic communication." Science **210**(4471): 801-803.

Shadlen, M. N. and J. I. Gold (2004). "The neurophysiology of decision-making as a window on cognition." The cognitive neurosciences **3**: 1229-1441.

Shannon, K. (2008). "The effect of rhythmic auditory stimulation on the gait parameters of patients with incomplete spinal cord injury: an exploratory pilot study." International Journal of Rehabilitation Research **31**(2): 155-157.

Shannon, K. (2010). "Neurologic music therapy: a scientific paradigm for clinical practice." Music and medicine **2**(2): 78-84.

Shennan, S. (2002). Genes, memes and human history. London, Thames and Hudson.

Shizgal, P. (1996). "The Janus faces of addiction." Behavioral and Brain Sciences **19**(4): 595-596.

Shofner, W. P. (2011). "Perception of the missing fundamental by chinchillas in the presence of low-pass masking noise." Journal of the Association for Research in Otolaryngology **12**(1): 101-112.

Shofner, W. P. and M. Chaney (2013). "Processing pitch in a nonhuman mammal (Chinchilla laniger)." Journal of Comparative Psychology **127**(2): 142.

Shove, P. and B. H. Repp (1995). "Musical motion and performance: theoretical and empirical perspectives." The practice of performance: Studies in musical interpretation: 55-83.

Simon, H. A., Ed. (1994). The bottleneck of attention: connecting thought with motivation. Integrative Views of Motivation, Cognition and Emotion Lincoln, University of Nebraska Press.

Simonov, P. V. e. (2013). The emotional brain: Physiology, neuroanatomy, psychology, and emotion, Springer Science & Business Media.

Simonyan, K. and B. Horwitz (2011). "Laryngeal motor cortex and control of speech in humans." The Neuroscientist **17**(2): 197-208.

Sinnott, J. M., C. H. Brown, et al. (1992). "Frequency and intensity discrimination in Mongolian gerbils, African monkeys and humans." Hearing research **59**(2): 205-212.

Sinnott, J. M., M. J. Owren, et al. (1987). "Auditory frequency discrimination in primates: Species differences (Cercopithecus, Macaca, Homo)." Journal of Comparative Psychology **101**(2): 126-131.

Sinnott, J. M., M. R. Petersen, et al. (1985). "Frequency and intensity discrimination in humans and monkeys." The Journal of the Acoustical Society of America **78**(6): 1977-1985.

Sjerps, M. J. and A. S. Meyer (2015). "Variation in dual-task performance reveals late initiation of speech planning in turn-taking." Cognition **136**: 304-324.

Skipper, J. I., H. C. Nusbaum, et al. (2005). "Listening to talking faces: motor cortical activation during speech perception." Neuroimage **25**(1): 76-89.

Slobin, M. (1993). Subcultural sounds: Micromusics of the West. Hanover, Wesleyan University Press.

Sloboda, J. (1991). "Music structure and emotional response: Some empirical findings." Psychology of Music **19**(19): 110-120.

Sloboda, J. (1992). Empirical studies of emotional response to music. Cognitive bases of musical communication. M. Riess-Jones and S. Holleran. Washington, DC, American Psychological Association**:** 33-46.

Sloboda, J. and P. Juslin (2001). Psychological perspectives on music and emotion. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press.

Sloboda, J. and S. O'Neill (2001). Emotions in everyday listening to music. Music and emotion: Theory and research. P. Juslin and J. Sloboda. Oxford, Oxford University Press**:** 415-430.

Small, C. (1987). "Music of the common tongue." London: John Calder.

Small, C. (1998). The meanings of performing and listening. Hanover, Weslyan University Press.

Smith, A. (1977/1993). Of the nature of that imitation which takes place in what are called the imitative arts. Essays on philosophical subjects. W. P. D. Wightman and J. C. Bryce, Liberty Fund, Indianapolis, IN (With Dugald
Stewart's account of Adam Smith, edited by IS Ross. General editors, DD Raphael and AS Skinner).
: 176-213.

Smith, A. D. (2014). "Models of language evolution and change." Wiley Interdisciplinary Reviews: Cognitive Science **5**(3): 281-293.

Smith, W. (1977). The Behavior of Communicating: An Ethological Approach. Harvard University Press, Cambridge, MA.

Sober, E. and D. Wilson (1998). Unto others: The evolution and psychology of unselfish behaviour. Cambridge, MA, Harvard University Press.

Sofianidis, G., V. Hatzitaki, et al. (2012). "Somatosensory driven interpersonal synchrony during rhythmic sway." Human Movement Science **31**(3): 553-566.

Song, X., M. S. Osmanski, et al. (2016). "Complex pitch perception mechanisms are shared by humans and a New World monkey." Proceedings of the National Academy of Sciences **113**(3): 781-786.

Songbirds, H. (2006). "The comparative psychology of absolute pitch." Comparative Cognition: Experimental Explorations of Animal Intelligence: Experimental Explorations of Animal Intelligence: **71**.

Spaulding, S. J., B. Barber, et al. (2013). "Cueing and gait improvement among people with Parkinson's disease: a meta-analysis." Archives of physical medicine and rehabilitation **94**(3): 562-570.

Speer, J. R. and P. U. Meeks (1985). "School children's perception of pitch in music." Psychomusicology: A Journal of Research in Music Cognition **5**(1-2): 49.

Spencer, H. (1857). "The origin and function of music." Essays, Scientific, Political, and Speculative **2**.

Spencer, H. (1857). "The origin and function of music." Fraser's Magazine **56**: 396-408.

Sperber, D. and L. Hirschfeld (1999). Culture, cognition and evolution. The MIT encyclopedia of congnitive sciences. R. Wilson and F. Keil. Cambridge, MA, MIT Press**:** cxi-cxxxii.

Sperber, D. and D. Wilson (1986). Relevance: Communication and cognition. Oxford, Blackwell.

Spiller, H. (2004). Gamelan: the traditional sounds of Indonesia, Abc-clio.

Spoor, F. and F. Zonneveld (1998). "Comparative review of the human body labyrinth." Yearbook of Physical Anthropology **41**: 211-251.

Sporns, O. (2006). "Small-world connectivity, motif composition, and complexity of fractal neuronal connections." Biosystems **85**(1): 55-64.

Sporns, O. and J. D. Zwi (2004). "The small world of the cerebral cortex." Neuroinformatics **2**(2): 145-162.

Spruijt, B., J. van Hooff, et al. (1992). "Ethology and neurobiology of grooming behavior." Physiological Review **72**(3): 825-852.

Sroufe, L. A. (2000). "Early relationships and the development of children." Infant Mental Health Journal **21**(1-2): 67-74.

Stahl, B., S. A. Kotz, et al. (2011). "Rhythm in disguise: why singing may not hold the key to recovery from aphasia." Brain **134**(10): 3083-3093.

Stainsby, T. and I. Cross (2009/2016). "The perception of pitch." The Oxford handbook of music psychology: 47-58.

Stebbins, W. C. (1973). "Hearing of old world monkeys (Cercopithecinae)." American Journal of Physical Anthropology **38**(2): 357-364.

Stefanics, G., B. Hangya, et al. (2010). "Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed." Journal of Neuroscience **30**(41): 13578-13585.

Steinbeis, N., S. Koelsch, et al. (2005). "Emotional processing of harmonic expectancy violations." Annals of the New York Academy of Sciences **1060**: 458-462.

Steinke, W., L. Cuddy, et al. (1997). "Dissociation of musical tonality and pitch memory from nonmusical cognitive abilities." Canadian Journal of Experimental Psychology **51**: 316-335.

Steinke, W., L. Cuddy, et al. (2001). "Dissociations among functional subsystems governing melody recognition after right-hemisphere damage." Cognitive Neuropsychology **18**(5): 411-437.

Stephan, K., M. Thaut, et al. (2002). "Conscious and subconscious sensorimotor synchronization—prefrontal cortex and the influence of awareness." Neuroimage **15**(2): 345-352.

Stephan, K., M. Thaut, et al. (2002). "Conscious and subconscious sensorimotor synchronization - prefrontal cortex and the influence of awareness." Neuroimage **15**(2): 345-352.

Steptoe, A. (2001). "Negative emotions in music making: The problem of performance anxiety." In P. N. Juslin & J. A. Sloboda (Eds.), Series in affective science. Music and emotion: Theory and research (pp. 291-307). Oxford University Press.

Stern, D. (1985). The interpersonal world of the infant: A view from psychoanalysis and developmental psychology. New York, Basic Books.

Stern, D. (1995). The motherhood constellation: A unified view of parent-infant psychotherapy. New York, Basic Books.

Stern, D. N. (1993). The role of feelings for an interpersonal self. Emory symposia in cognition, 5. The perceived self: Ecological and interpersonal sources of self-knowledge. U. Neisser. New York, Cambridge University Press**:** 205-215.

Stern, D. N. (2004). The present moment in psychotherapy and everyday life (norton series on interpersonal neurobiology), WW Norton & Company.

Stern, D. N. (2009). The first relationship, Harvard University Press.

Stern, D. N. (2010). Forms of vitality: Exploring dynamic experience in psychology, the arts, psychotherapy, and development, Oxford University Press.

Stevens, C. and T. Byron (2009/16). "Universals in music processing." Oxford handbook of music psychology: 14-23.

Stevens, L. T. (1886). "On the time-sense." Mind **11**(43): 393-404.

Stivers, T., N. J. Enfield, et al. (2009). "Universals and cultural variation in turn-taking in conversation." Proceedings of the National Academy of Sciences **106**(26): 10587-10592.

Stobart, H. and I. Cross (2000). "The Andean anacrusis? Rhythmic structure and perception in Easter songs of Northern Potosi, Bolivia." British Journal of Ethnomusicology **9**(2): 63-94.

Stoeger, A. S. and P. Manger (2014). "Vocal learning in elephants: Neural bases and adaptive context." Current opinion in neurobiology **28**: 101-107.

Stoeger, A. S., D. Mietchen, et al. (2012). "An Asian elephant imitates human speech." Current Biology **22**(22): 2144-2148.

Stokes, M., Ed. (1994). Ethnicity, identity and music. Oxford, Berg.

Stoklasa, J., C. Liebermann, et al. (2012). Timing and synchronization of professional musicians: A comparison between orchestral brass and string players. 12th international conference on music perception and cognition, Thessaloniki, Greece.

Storr, A. (1992). Music and the mind. London, Harper-Collins.

Sugimoto, T., H. Kobayashi, et al. (2010). "Preference for consonant music over dissonant music by an infant chimpanzee." Primates **51**(1): 7.

Sumbre, G., A. Muto, et al. (2008). "Entrained rhythmic activities of neuronal ensembles as perceptual memory of time interval." Nature **456**(7218): 102.

Sundberg, J. (1987). "The science of the singing voice". DeKalb, IL, Northern Illinois Press."

Sundberg, J. and T. D. Rossing (1990). "The science of singing voice." The Journal of the Acoustical Society of America **87**(1): 462-463.

Surian, L., S. Caldi, et al. (2007). "Attribution of beliefs by 13-month-old infants." Psychological science **18**(7): 580-586.

Sutton, D., C. Larson, et al. (1973). "Vocalization in rhesus monkeys: Conditionability." Brain research **52**: 225-231.

Sutton, D., R. Trachy, et al. (1985). "Discriminative phonation in macaques: Effects of anterior mesial cortex damage." Experimental Brain Research **59**(2): 410-413.

Svanibor, P. (1997). Music, Politics and War: Views from Croatia. Zagreb, Institute of Ethnology and Folklore Research.

Swaminathan, S. and E. G. Schellenberg (2015). "Current emotion research in music psychology." Emotion review **7**(2): 189-197.

Swann Jr, W. B., J. Jetten, et al. (2012). "When group membership gets personal: a theory of identity fusion." Psychological review **119**(3): 441.

Szelag, E. (1997). "Temporal integration of the brain as studied with the metronome paradigm." Time, temporality, now. Experiencing time and concepts of time in an interdisciplinary perspective: 121-131.

Szelag, E., N. von Steinbüchel, et al. (1995). "Temporal constraints in processing of nonverbal rhythmic patterns." Acta Neurobiologiae Experimentalis **56**(1): 215-225.

Taglialatela, J. P., S. Savage-Rumbaugh, et al. (2003). "Vocal production by a language-competent Pan paniscus." International Journal of Primatology **24**(1): 1-17.

Takahashi, D. Y., D. Z. Narayanan, et al. (2013). "Coupled oscillator dynamics of vocal turn-taking in monkeys." Current Biology **23**(21): 2162-2168.

Takahashi, N., M. Kawamura, et al. (1992). "Pure word deafness due to left hemisphere damage." Cortex **28**(2): 295-303.

Takano, K. and Y. Miyake (2007). "Two types of phase correction mechanism involved in synchronized tapping." Neuroscience letters **417**(2): 196-200.

Tallerman, M. (2005). Language origins: Perspectives on evolution, Oxford University Press, USA.

Tallerman, M. (2007). "Did our ancestors speak a holistic protolanguage?" Lingua **117**(3): 579-604.

Tallerman, M. (2008). "Holophrastic protolanguage: Planning, processing, storage, and retrieval." Interaction Studies **9**(1): 84-99.

Tan, Y. T., G. E. McPherson, et al. (2014). "The genetic basis of music ability." Frontiers in psychology **5**.

Tanabe, H. C., M. Kato, et al. (2005). "The sensorimotor transformation of cross-modal spatial information in the anterior intraparietal sulcus as revealed by functional MRI." Cognitive Brain Research **22**(3): 385-396.

Tang, Y.-Y., Y. Ma, et al. (2007). "Short-term meditation training improves attention and self-regulation." Proceedings of the National Academy of Sciences **104**(43): 17152-17156.

Tarr, B., J. Launay, et al. (2015). "Synchrony and exertion during dance independently raise pain threshold and encourage social bonding." Biology letters **11**(10): 20150767.

Tarr, B., J. Launay, et al. (2016). "Silent disco: dancing in synchrony leads to elevated pain thresholds and social closeness." Evolution and Human Behavior **37**(5): 343-349.

Teki, S., M. Grube, et al. (2011). "Distinct neural substrates of duration-based and beat-based auditory timing." Journal of Neuroscience **31**(10): 3805-3812.

ten Cate, C., M. Spierings, et al. (2016). "Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species." Frontiers in psychology **7**.

Terhardt, E. (1984). "The concept of musical consonance: A link between music and psychoacoustics." Music Perception: An Interdisciplinary Journal **1**(3): 276-295.

Thaut, M. (1990). Neuropschological processes in music perception and their relevance in music therapy. In Music therapy in the treatment of adults with mental disorders. R. Unkefer (ed.). New York, Schirmer Books.

Thaut, M. (2002). "Psychomusical foundations of music therapy: neuropsychological processes in music perception and their relevance in music therapy." Music therapy in the treatment of adults with mental disorders–theoretical bases and clinical information: 12-24.

Thaut, M., G. Kenyon, et al. (2002). "Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients." Neuropsychologia **40**(7): 1073-1081.

Thaut, M., G. Kenyon, et al. (1999). "The connection between rhythmicity and brain function." IEEE Engineering in Medicine and Biology Magazine **18**(2): 101-108.

Thaut, M., A. Leins, et al. (2007). "Rhythmic auditor y stimulation improves gait more than NDT/Bobath training in near-ambulatory patients early poststroke: a single-blind, randomized trial." Neurorehabilitation and neural repair **21**(5): 455-459.

Thaut, M., G. McIntosh, et al. (1997). "Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation." Journal of Neurological Sciences **15**(1): 207-212.

Thaut, M., J. Rathbun, et al. (1997). "Music versus metronome timekeeper in a rhythmic motor task." International Journal of Arts Medicine **5**: 4-12.

Thaut, M., B. Tian, et al. (1998). "Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment." Human Movement Science **17**(6): 839-863.

Thaut, M. H. (2002). "Toward a cognition-affect model in neuropsychiatric music therapy." Music Therapy in the Treatment of Adults with Mental Disorders: Theoretical Bases and Clinical Interventions, eds RF Unkefer and MT Thaut (Gilsum, NH: Barcelona Publishers): 86-103.

Thaut, M. H. (2005). Rhythm, music, and the brain: Scientific foundations and clinical applications, Routledge.

Thaut, M. H. (2015). "The discovery of human auditory–motor entrainment and its role in the development of neurologic music therapy." Progress in brain research **217**: 253-266.

Thaut, M. H. and M. Abiru (2010). "Rhythmic auditory stimulation in rehabilitation of movement disorders: a review of current research." Music Perception: An Interdisciplinary Journal **27**(4): 263-269.

Thaut, M. H., G. C. McIntosh, et al. (2014). "Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system." Frontiers in psychology **5**.

Thaut, M. H., G. C. McIntosh, et al. (1992). "Effect of rhythmic auditory cuing on temporal stride parameters and EMG patterns in normal gait." Journal of Neurologic Rehabilitation **6**(4): 185-190.

Thaut, M. H., G. C. McIntosh, et al. (1993). "Effect of rhythmic auditory cuing on temporal stride parameters and EMG. Patterns in hemiparetic gait of stroke patients." Journal of Neurologic Rehabilitation **7**(1): 9-16.

Thaut, M. H., G. C. McIntosh, et al. (1996). "Rhythmic auditory stimulation in gait training for Parkinson's disease patients." Movement disorders **11**(2): 193-200.

Thaut, M. H., K. W. McIntosh, et al. (2001). "Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease." Functional neurology **16**(2): 163-172.

Thaut, M. H., K. M. Stephan, et al. (2009). "Distinct cortico-cerebellar activations in rhythmic auditory motor synchronization." Cortex **45**(1): 44-53.

Thayer, R. (1978). "Toward a psychological theory of multidimensional activation (arousal)." Motivation and Emotion **2**: 1-34.

Thayer, R. (1986). Activation (arousal): The shift from a single to a multidimensional perspective. The biological basis of personality and behaviour. J. Strelau, F. Farley and A. Gale. London, Hemisphere. **1:** 115-127.

Thayer, R. (1989). The biopsychology of mood and arousal. Oxford Oxford University Press.

Thayer, R., J. Newman, et al. (1994). "The self-regulation of mood: Strategies for changing a bad mood, raising energy, and reducing tension." Journal of Personality and Social Psychology **67**(5): 910-925.

Thompson, W. and L. Balkwill (2006). "Decoding speech prosody in five languages." Semiotica **158**(1): 407-424.

Thompson, W. and B. Robitaille (1992). "Can composers express emotions through music?" Empirical Studies of the Arts **10**: 79-89.

Thompson, W. F. (2013). "Intervals and scales." The psychology of music **3**: 107-140.

Thompson, W. F., M. M. Marin, et al. (2012). "Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis." Proceedings of the National Academy of Sciences **109**(46): 19027-19032.

Thompson, W. F. and L. Quinto (2011). "Music and emotion: Psychological considerations." The Aesthetic Mind: Philosophy and Psychology: 357-375.

Thorpe, L. and S. Trehub (1989). "Duration illusion and auditory grouping in infancy." Developmental Psychology **25**: 122-127.

Thorpe, W. H., J. Hall-Craggs, et al. (1972). "Duetting and antiphonal song in birds: its extent and significance." Behaviour. Supplement(18): III-197.

Tian, X. and D. Poeppel (2010). "Mental imagery of speech and movement implicates the dynamics of internal forward models." Frontiers in psychology **1**.

Tillmann, B., J. Bharucha, et al. (2000). "Implicit learing of tonality: A self-organizing approach." Psychological Review **107**(4): 885-913.

Tillmann, B., D. Burnham, et al. (2011). "Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages." Frontiers in psychology **2**.

Tillmann, B., Y. Lévêque, et al. (2016). "Impaired short-term memory for pitch in congenital amusia." Brain research **1640**: 251-263.

Tillmann, B., E. Rusconi, et al. (2011). "Fine-grained pitch processing of music and speech in congenital amusia." The Journal of the Acoustical Society of America **130**(6): 4089-4096.

Tillmann, B., K. Schulze, et al. (2009). "Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds." Brain and cognition **71**(3): 259-264.

Tinbergen, N. (1952). "" Derived" activities; their causation, biological significance, origin, and emancipation during evolution." The Quarterly review of biology **27**(1): 1-32.

Tinbergen, N. (1963). "On aims and methods of ethology." Ethology **20**(4): 410-433.

Tincoff, R., M. Hauser, et al. (2005). "The role of speech rhythm in language discrimination: further tests with a non-human primate." Developmental science **8**(1): 26-35.

Titon, J., Ed. (1996). Worlds of music: an introduction to the music of the world's people (3rd edition). New York, Schirmer.

Titon, J. T. (2015). Worlds of Music: an Introduction to the Music of the World’s Peoples, Nelson Education.

Titze, I. R. (1994). Principles of voice production, Allyn & Bacon.

Toda, S., A. Fogel, et al. (1990). "Maternal speech to three-month-old infants in the United States and Japan." Journal of Child Language **17**: 279-294.

Todd, N. P. and C. S. Lee (2015). "The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives." Frontiers in human neuroscience **9**.

Todd, N. P. M. (1995). "The kinematics of musical expression." The Journal of the Acoustical Society of America **97**(3): 1940-1949.

Todd, N. P. M. (1999). "Motion in music: A neurobiological perspective." Music Perception: An Interdisciplinary Journal **17**(1): 115-126.

Tolbert, E. (2001). "Music and meaning: An evolutionary story." Psychology of Music **29**: 89-94.

Tolbert, E. (2008). "The Karelian lament." In M. Clayton (ed.) Music, words and voice: a reader: 202-212.

Tomasello, M. (1999). "The human adaptation for culture." Annual Review of Anthropology **28**: 509-529.

Tomasello, M. (2007). inThe Gestural Communication of Apes and Monkeys ed. Call J., Tomasello M, Lawrence Erlbaum Associates.

Tomasello, M. (2009). The cultural origins of human cognition, Harvard University Press.

Tomasello, M. (2009). Why we cooperate, MIT Press.

Tomasello, M. (2010). Origins of human communication, MIT Press.

Tomasello, M. and J. Call (1997). Primate cognition, Oxford University Press, USA.

Tomasello, M., J. Call, et al. (1998). "Five primate species follow the visual gaze of conspecifics." Animal Behaviour **55**(4): 1063-1069.

Tomasello, M., J. Call, et al. (2003). "Chimpanzees understand psychological states–the question is which ones and to what extent." Trends in cognitive sciences **7**(4): 153-156.

Tomasello, M., M. Carpenter, et al. (2005). "Understanding and sharing intentions: The origins of cultural cognition." Behavioural and Brain Sciences **28**(5): 675-691.

Tomasello, M. and K. Haberl (2003). "Understanding attention: 12-and 18-month-olds know what is new for other persons." Developmental Psychology **39**(5): 906.

Tomasello, M., B. Hare, et al. (1999). "Chimpanzees, Pan troglodytes, follow gaze direction geometrically." Animal Behaviour **58**(4): 769-777.

Tomasello, M., B. Hare, et al. (2001). "The ontogeny of gaze following in chimpanzees, Pan troglodytes, and rhesus macaques, Macaca mulatta." Animal Behaviour **61**(2): 335-343.

Tomlinson, G. (2015). A million years of music: the emergence of human modernity, MIT Press.

Tomlinson, G. (2018). Culture and the Course of Human Evolution. University of Chicago Press.

Tomlinson, R. W. and D. W. Schwarz (1988). "Perception of the missing fundamental in nonhuman primates." The Journal of the Acoustical Society of America **84**(2): 560-565.

Tomonaga, M., M. Tanaka, et al. (2004). "Development of social cognition in infant chimpanzees (Pan troglodytes): Face recognition, smiling, gaze, and the lack of triadic interactions." Japanese Psychological Research **46**(3): 227-235.

Tooby, J. and L. Cosmides (1992). The psychological foundations of culture. The adapted mind: Evolutionary psychology and the generation of culture. J. Barkow, L. Cosmides and J. Tooby. Oxford, Oxford University Press**:** 19-136.

Toro, J. M. and P. Crespo-Bojorque (2017). "Consonance Processing in the Absence of Relevant Experience: Evidence from Nonhuman Animals." Comparative Cognition & Behavior Reviews **12**.

Toro, J. M., J. B. Trobalon, et al. (2003). "The use of prosodic cues in language discrimination tasks by rats." Animal cognition **6**(2): 131-136.

Torre, K. and R. Balasubramaniam (2009). "Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements." Experimental brain research **199**(2): 157-166.

Townsend, S. W., B. D. Charlton, et al. (2014). "Acoustic cues to identity and predator context in meerkat barks." Animal Behaviour **94**: 143-149.

Toyomura, A., S. Koyama, et al. (2007). "Neural correlates of auditory feedback control in human." Neuroscience **146**(2): 499-503.

Trainor, L. (1996). "Infant preferences for infant-directed versus non-infant directed play songs and lullabies." Infant Behaviour and Development **19**: 83-92.

Trainor, L. (1997). "Effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals." Journal of Experimental Psychology: Human Perception and Performance **23**: 1427-1438.

Trainor, L. (2005). "Are there critical periods for music development?" Developmental Psychobiology **46**: 262-278.

Trainor, L. and B. Adams (2000). "Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns." Perception and Psychophysics **2**: 333-340.

Trainor, L., C. Austin, et al. (2000). "Is infant-directed speech prosody a result of the vocal expression of emotion?" Psychological Science **11**: 188-195.

Trainor, L., E. Clark, et al. (1997). "The acoustic basis of infant preferences for infant-directed singing." Infant Behaviour and Development **20**: 383-396.

Trainor, L. and S. Trehub (1992). "A comparison of infants' and adults' sensitivity to Western musical structure." Journal of Experimental Psychology: Human Perception and Performance **18**: 394-402.

Trainor, L. and S. Trehub (1993). "What mediates infants' and adults' superior processing of the major over the augmented triad?" Music Perception **11**: 185-196.

Trainor, L. and S. Trehub (1994). "Key membership and implied harmony in Western tonal music: Developmental perspectives." Perception and Psychophysics **56**: 125-132.

Trainor, L., L. Wu, et al. (2004). "Long-term memory for music: Infants remember tempo and timbre." Developmental Science **7**: 289-296.

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

Trainor, L. J. and K. A. Corrigall (2010). Music acquisition and effects of musical experience. Music perception, Springer**:** 89-127.

Trainor, L. J., X. Gao, et al. (2009). "The primal role of the vestibular system in determining musical rhythm." Cortex **45**(1): 35-43.

Trainor, L. J. and B. M. Heinmiller (1998). "The development of evaluative responses to music:: infants prefer to listen to consonance over dissonance." Infant Behavior and Development **21**(1): 77-88.

Trainor, L. J., C. Marie, et al. (2012). "Becoming musically enculturated: effects of music classes for infants on brain and behavior." Annals of the New York Academy of Sciences **1252**(1): 129-138.

Trainor, L. J., C. D. Tsang, et al. (2002). "Preference for sensory consonance in 2-and 4-month-old infants." Music Perception: An Interdisciplinary Journal **20**(2): 187-194.

Trainor, L. J. and A. Unrau (2009). "Extracting the beat: An experience-dependent complex integration of multisensory information involving multiple levels of the nervous system."

Tramo, M. J., P. A. Cariani, et al. (2001). "Neurobiological foundations for the theory of harmony in Western tonal music." Annals of the New York Academy of sciences **930**(1): 92-116.

Trehub, S. (2000). Human Processing Predispositions. The origins of music. N. Wallin, B. Merker and S. Brown. Cambridge, MA, MIT Press.

Trehub, S. (2003). "Commentary on Saffron and Griegentrog (2001): absolute pitch and relative pitch processing in tone learning tasks." Developmental Science **6**: 46-47.

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

Trehub, S. (2003). Musical predispositions in infancy: An update. The cognitive neuroscience of music. I. Peretz and R. Zatorre. Oxford, Oxford University Press**:** 3-20.

Trehub, S., D. Bull, et al. (1984). "Infants' perception of melodies: The role of melodic contour." Child Develoment **55**: 821-830.

Trehub, S., L. Endman, et al. (1990). "Infants' perception of timbre: classification of complex tones by spectral structure." Journal of Experimental Child Psychology **49**: 300-313.

Trehub, S. and E. Hannon (2006). "Infant music perception: domain-general or domain-specific mechanisms?" Cognition **100**(1): 73-99.

Trehub, S., E. Schellenberg, et al. (1999). "Infants' and adults' perception of scale structure." Journal of Experimental Psychology: Human Perception and Performance **25**: 965-975.

Trehub, S., J. Schneider, et al. (1995). "Gap detection in infants, children, and adults." Journal of the Acoustical Society of America **98**: 2532-2541.

Trehub, S. and L. Thorpe (1989). "Infants' perception of rhythm: Categorization of auditory sequences by temporal structure." Canadian Journal of Psychology **43**: 217-220.

Trehub, S., L. Thorpe, et al. (1987). "Organisational processes in infants' perception of auditory patterns." Child Develoment **58**: 741-749.

Trehub, S., L. Thorpe, et al. (1990). "Infants' perception of good and bad melodies." Psychomusicology **9**: 5-19.

Trehub, S., A. Unyk, et al. (1993a). "Adults identify infant-directed music across cultures." Infant Behaviour and Development **16**: 193-211.

Trehub, S., A. Unyk, et al. (1993b). "Maternal singing in cross-cultural perspective." Infant Behaviour and Development **16**: 285-295.

Trehub, S. E. (1990). The perception of musical patterns by human infants: The provision of similar patterns by their parents. Comparative perception. M. A. B. a. W. C. Stebbins. New York: Wiley. **1: Mechanisms:** 429-459.

Trehub, S. E., J. Becker, et al. (2015). "Cross-cultural perspectives on music and musicality." Phil. Trans. R. Soc. B **370**(1664): 20140096.

Trehub, S. E., A. J. Cohen, et al. (1986). "Development of the perception of musical relations: Semitone and diatonic structure." Journal of experimental psychology: Human perception and performance **12**(3): 295.

Trehub, S. E. and L. Trainor (1998). "Singing to infants: Lullabies and play songs." Advances in infancy research **12**: 43-78.

Trehub, S. E., A. M. Unyk, et al. (1997). "Mothers' and fathers' singing to infants." Developmental psychology **33**(3): 500.

Treisman, M. (1963). "Temporal discrimination and the indifference interval: Implications for a model of the" internal clock"." Psychological Monographs: General and Applied **77**(13): 1.

Tremblay-Champoux, A., S. Dalla Bella, et al. (2010). "Singing proficiency in congenital amusia: imitation helps." Cognitive neuropsychology **27**(6): 463-476.

Trevarthen, C. (1993). "The self born in intersubjectivity: The psychology of an infant communicating." In U. Neisser (Ed.), Emory symposia in cognition, 5. The perceived self: Ecological and interpersonal sources of self-knowledge (pp. 121-173). New York, NY, US: Cambridge University Press.

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

Trevarthen, C. (1999). "Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication." Musicae Scientiae **3**(1_suppl): 155-215.

Trevarthen, C. (2002). "Origins of musical identity: Evidence from infancy for musical social awareness." In MacDonald, Hargreaves & Miell (Eds.) Musical identities: 21-38. OUP.

Trevarthen, C. (2005). "Action and emotion in development of cultural intelligence: Why infants have feelings like ours." Emotional development: 61-91.

Trevarthen, C. (2012). "The generation of human meaning: How shared experience grows in infancy." Joint Attention: new developments in psychology, philosophy of mind, and social neuroscience: 73-113.

Trevarthen, C. (2015). "Infant Semiosis: The psycho-biology of action and shared experience from birth." Cognitive Development **36**: 130-141.

Trevarthen, C. (2016). From the Intrinsic Motive Pulse of Infant Actions to the Life Time of Cultural Meanings. Philosophy and Psychology of Time, Springer**:** 225-265.

Trevarthen, C. and K. Aitken (1995). "Brain development, infant communication and empathy disorders: Intrinsic factors in child mental health." Development and Psychopathology **6**: 597-633.

Trevarthen, C. and K. Aitken (2001). "Infant intersubjectivity: research, theory, and clinical applications." Journal of Child Psychology and Psychiatry and Allied Disciplines **42**(1): 3-48.

Trevarthen, C., J. Delafield-Butt, et al. (2011). "Psychobiology of Musical Gesture: Innate Rhythm, Harmony and Melody." New perspectives on music and gesture: 11-43.

Trevarthen, C. and J. T. Delafield-Butt (2013). "Autism as a developmental disorder in intentional movement and affective engagement." Frontiers in Integrative Neuroscience **7**: 49.

Tronick, E. Z. (1989). "Emotions and emotional communication in infants." American psychologist **44**(2): 112.

Truslit, A. (2015). Gestaltung und Bewegung in der Musik: Ein tönendes Buch vom musikalischen Vortrag und seinem bewegungserlebten Gestalten und Hören-Reprint der Ausgabe Berlin 1938, Wißner-Verlag.

Tschacher, W., G. M. Rees, et al. (2014). "Nonverbal synchrony and affect in dyadic interactions." Frontiers in psychology **5**: 1323.

Tunçgenç, B. and E. Cohen (2016). "Movement synchrony forges social bonds across group divides." Frontiers in psychology **7**: 782.

Tunçgenç, B. and E. Cohen (2018). "Interpersonal movement synchrony facilitates pro-social behavior in children's peer-play." Developmental science **21**(1).

Tunçgenç, B., E. Cohen, et al. (2015). "Rock with me: The role of movement synchrony in infants' social and nonsocial choices." Child development **86**(3): 976-984.

Tuomela, R. (1995). The importance of us: A philosophical study of basic social notions, Stanford University Press.

Turgeon, M., A. M. Wing, et al. (2011). "Timing and aging: Slowing of fastest regular tapping rate with preserved timing error detection and correction." Psychology and aging **26**(1): 150.

Turino, T. (1991). "The history of a Peruvian panpipe style and the politics of interpretation." Ethnomusicology and modern music history: 121-138.

Turk, A. and S. Shattuck-Hufnagel (2013). "What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong." Laboratory Phonology **4**(1): 93-118.

Turk, I., Ed. (1997). Mousterian "bone flute" and other finds from Divje babe I cave site in Slovenia. Ljubljana, Institut za archaeologijo, Znanstvernorazixkovalni Center Sazu.

Turk, I., J. Dirjec, et al. (1997). Description and explanation of the origin of the suspected bone flute. Mousterian "bone flute" and other finds from Divje babe I cave site in Slovenia.. Ljubljana, Institut za archaeologijo, Znanstvernorazixkovalni Center Sazu.

Turk, I. and B. Kavur (1997). Paleolithic bone flutes - comparable material. Mousterian "bone flute" and other finds from Divje babe I cave site in Slovenia. Ljublijana, Institut za archaeologijo, Znanstvernorazixkovalni Center Sazu**:** 179-184.

Turnbull, C. (1978). "Rethinking the Ik: a functional non-social system." Extinction and Survival in Human Populations. Columbia University Press, New York: 49-75.

Turner, V. (1974). Dramas, fields and metaphors: Symbolic action in human society. Ithaca, NY, Cornell University Press.

Turner, V. (1991). The ritual process. Structure and anti-structure. Ithaca, NY, Cornell University Press.

Turner, V. (2017). The ritual process: Structure and anti-structure, Routledge.

Tylor (1871). Primitive culture. London, Murray.

Ukkola, L. T., P. Onkamo, et al. (2009). "Musical aptitude is associated with AVPR1A-haplotypes." PloS one **4**(5): e5534.

Ullén, F., H. Forssberg, et al. (2003). "Neural networks for the coordination of the hands in time." Journal of neurophysiology **89**(2): 1126-1135.

Unkefer, R. and M. H. Thaut (2005). Music Therapy in the Treatment of Adults With Mental Disorders: Theoretical Bases and Clinical Interventions, Barcelona Publishers.

Unyk, A., S. Trehub, et al. (1992). "Lullabies and simplicity: A cross-cultural perspective." Psychology of Music **20**: 15-28.

Vaestfjall, D. (2002). "A review of the musical mood induction procedure." Musicae Scientiae **Special Issue 2001-2002**: 173-211.

Vaitl, D., W. Vehrs, et al. (1993). Prompts - leitmotif - emotion: play it again, Richard Wagner. The structure of emotion: Psychophysiological, cognitive, and clinical aspects. N. Birnbaumer and A. Oehman. Seattle, WA, Hogrefe and Huber**:** 169-189.

Valdesolo, P. and D. DeSteno (2011). "Synchrony and the social tuning of compassion." Emotion **11**(2): 262.

Valdesolo, P., J. Ouyang, et al. (2010). "The rhythm of joint action: Synchrony promotes cooperative ability." Journal of Experimental Social Psychology **46**(4): 693-695.

Valentine, C. (1962). "Musical intervals and attitudes to music." In C. W. Valentine (Ed.), The experimental psychology of beauty (pp. 196 –227). London, UK: Methuen & Co. Ltd.

Van Baaren, R. B., R. W. Holland, et al. (2004). "Mimicry and prosocial behavior." Psychological science **15**(1): 71-74.

Van Baaren, R. B., R. W. Holland, et al. (2003). "Mimicry for money: Behavioral consequences of imitation." Journal of Experimental Social Psychology **39**(4): 393-398.

van der Schyff, D. (2013). "Emotion, embodied mind and the therapeutic aspects of musical experience in everyday life." Approaches: Music Therapy and Special Music Education **5**(1): 50-58.

van der Schyff, D. and A. Schiavio (2017). "Evolutionary Musicology Meets Embodied Cognition: Biocultural Coevolution and the Enactive Origins of Human Musicality." Frontiers in Neuroscience **11**: 519.

van der Schyff, D. and A. Schiavio (2017). "The future of musical emotions." Frontiers in psychology **8**: 988.

van der Steen, M. M. and P. E. Keller (2013). "The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization." Frontiers in human neuroscience **7**.

Van Gennep, A. and G. L. Caffee (1960). The Rites of Passage. Translated by Monika B. Vizedomand and Gabrielle L. Caffee. Routledge & Kegan Paul.

Van Puyvelde, M., G. Loots, et al. (2015). "A cross-cultural comparison of tonal synchrony and pitch imitation in the vocal dialogs of Belgian Flemish-speaking and Mexican Spanish-speaking mother–infant dyads." Infant behavior and development **40**: 41-53.

Van Puyvelde, M., G. Loots, et al. (2014). "Do You Hear the Same? Cardiorespiratory responses between mothers and infants during tonal and atonal music." PloS one **9**(9): e106920.

Van Puyvelde, M., G. Loots, et al. (2013). "The interplay between tonal synchrony and social engagement in mother–infant interaction." Infancy **18**(5): 849-872.

Van Puyvelde, M., P. Vanfleteren, et al. (2010). "Tonal synchrony in mother–infant interaction based on harmonic and pentatonic series." Infant behavior and development **33**(4): 387.

van Ulzen, N. R., C. J. Lamoth, et al. (2010). "Stability and variability of acoustically specified coordination patterns while walking side-by-side on a treadmill: Does the seagull effect hold?" Neuroscience letters **474**(2): 79-83.
VanderHorst, V., E. Terasawa, et al. (2001). "Monosynaptic projections from the nucleus retroambiguus region to laryngeal motoneurons in the rhesus monkey." Neuroscience **107**(1): 117-125.

Vaneechouette, M. and J. Skoyles (1998). "The memtic origin of language: modern humans as musical primates." Journal of Memetics - Evolutionary Models of Information Transmission **2**.
Varela, F. J., E. Thompson, et al. (2017). The embodied mind: Cognitive science and human experience, MIT press.

Vassilakis, P. N. (2005). "Auditory roughness as a means of musical expression." Selected Reports in Ethnomusicology **12**(119-144): 122.

Velasco, M. J. and E. W. Large (2011). "Pulse detection in syncopated rhythms using neural oscillators." Pulse **1**(2): 3-4.

Vervoort, V., R. Smith, et al. (2002). "Genomic rearrangements of EYA1 account for a large fraction of families with BOR syndrome." European Journal of Human Genetics **10**(11): 757-766.

Visco-Comandini, F., S. Ferrari-Toniolo, et al. (2015). "Do non-human primates cooperate? Evidences of motor coordination during a joint action task in macaque monkeys." cortex **70**: 115-127.

Vitouch, O., Ed. (2007). Musicae Scientiae, special issue on Music and Evolution.

Vohs, K. D. and R. F. Baumeister (2016). Handbook of self-regulation: Research, theory, and applications, Guilford Publications.

Volkova, A. (2004). Long-term memory for absolute pitch level of songs in infancy. Missisauga, Ontario, Canada, University of Toronto. **MA**.

Von Helmholtz, H. (2013). Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik, Springer-Verlag.

von Steinbüchel, N., M. Wittmann, et al. (1999). "Temporal constraints of perceiving, generating, and integrating information: Clinical indications." Restorative Neurology and Neuroscience **14**(2-3): 167-182.

von Zimmermann, J. and D. C. Richardson (2016). "Verbal Synchrony and Action Dynamics in Large Groups." Frontiers in psychology **7**: 2034.

Vorberg, D. and R. Hambuch (1984). "Timing of Two-Handed Rhythmic Performance." Annals of the New York Academy of sciences **423**(1): 390-406.

Vorberg, D. and H.-H. Schulze (2002). "Linear phase-correction in synchronization: Predictions, parameter estimation, and simulations." Journal of Mathematical Psychology **46**(1): 56-87.

Vorberg, D. and A. Wing (1996). "Modeling variability and dependence in timing." Handbook of perception and action **2**: 181-262.

Vos, P. G., J. Mates, et al. (1995). "The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies." The Quarterly journal of experimental psychology **48**(4): 1024-1040.

Vuvan, D., S. Paquette, et al. (2017). "The Montreal Protocol for Identification of Amusia." Behavior Research Methods: 1-11.

Vuvan, D. T., M. Nunes-Silva, et al. (2015). "Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia." cortex **69**: 186-200.

Wacewicz, S. (2014). "The narrow faculty of language: What is it, who has it, and how is it defined?". Theoria et Historia Scientiarum, **9**: 217–229.

Wacewicz, S. and P. Żywiczyński (2015). "Language evolution: Why Hockett's design features are a non-starter." Biosemiotics **8**(1): 29-46.

Wachsmann, K. P. (1971). "Universal perspectives in music." Ethnomusicology **15**(3): 381-384.

Wade, M. (1978). "A critical review of the models of group selection." Quarterly review of Biology **53**: 101-114.

Walker, A. (1993). Perspective on the Nariokotome discovery. The Nariokotome Home erectus skeleton. A. Walker and R. Leakey. Cambridge, MA, Harvard University Press**:** 411-432.

Waller, S. M., M. Harris-Love, et al. (2006). "Temporal coordination of the arms during bilateral simultaneous and sequential movements in patients with chronic hemiparesis." Experimental brain research **168**(3): 450-454.

Wallin, N., B. Merker, et al., Eds. (2000). The origins of music. Cambridge MA, MIT Press.

Wambaugh, J. L. and A. L. Martinez (2000). "Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech." Aphasiology **14**(8): 851-871.

Wang, T.-H., Y.-C. Peng, et al. (2013). "A home-based program using patterned sensory enhancement improves resistance exercise effects for children with cerebral palsy: a randomized controlled trial." Neurorehabilitation and neural repair **27**(8): 684-694.

Wang, X. (2013). "The harmonic organization of auditory cortex." Frontiers in systems neuroscience **7**.

Warren, J., A. Jennings, et al. (2005). "Analysis of the spectral envelope of sounds by the human brain." Neuroimage **24**(4): 1052-1057.

Warren, R. M. (1993). "Perception of acoustic sequences: Global integration versus temporal resolution." In S. McAdams & E. Bigand (Eds.), Oxford science publications. Thinking in sound: The cognitive psychology of human audition (pp. 37-68). New York, NY, US: Clarendon Press/Oxford University Press.

Watanabee, J. and B. Smuts (1999). "Explaining ritual without explaining it away." American Anthropologist **101**: 98-112.

Waterman, M. (1996). "Emotional responses to music: Implicit and explicit effects in listeners and performers." Psychology of Music **24**: 53-67.

Watson, K. (1942). "The nature and measurement of musical meanings." Psychological Monographs **54**: 1-43.

Watts, C., R. Moore, et al. (2005). "The relationship between vocal pitch-matching skills and pitch discrimination skills in untrained accurate and inaccurate singers." Journal of Voice **19**(4): 534-543.

Webb, J. (1972). Interview synchrony: an investigation into two speech rate measures. Studies in dyadic communication. A. Siegman and B. Pope. New York, Pergamon Press.

Wedin, L. (1969). "Dimension analysis of emotional expression in music." Swedish Journal of Musicology **51**: 119-140.

Weinstein, D., J. Launay, et al. (2016). "Group music performance causes elevated pain thresholds and social bonding in small and large groups of singers." Evolution and human behavior: official journal of the Human Behavior and Evolution Society **37**(2): 152.

Weinstein, D., J. Launay, et al. (2016). "Singing and social bonding: Changes in connectivity and pain threshold as a function of group size." Evolution and Human Behavior **37**(2): 152-158.

Weisman, R. G., M. G. Njegovan, et al. (2004). "A behavior analysis of absolute pitch: sex, experience, and species." Behavioural processes **66**(3): 289-307.

Wellman, H. M., D. Cross, et al. (2001). "Meta-analysis of theory-of-mind development: the truth about false belief." Child development **72**(3): 655-684.

Wells, A. (1990). "Popular music: Emotional use and management." Journal of Popular Culture **24**: 105-117.

Wen, N. J., P. A. Herrmann, et al. (2016). "Ritual increases children’s affiliation with in-group members." Evolution and Human Behavior **37**(1): 54-60.

Wermke, K. and W. Mende (2009). "Musical elements in human infants’ cries: in the beginning is the melody." Musicae Scientiae **13**(2_suppl): 151-175.

Werner, D. (1984). Amazon journey; An anthropologist's year among Brazil's Mekranoti Indians. New York, Simon & Shuster.

Werner, L. (1992). Interpreting developmental psycholacoustics. Develomental psychoacoustics. L. Werner and E. Rubel. Washington, DC, American Psychological Assocation.

Werner, L., G. Marean, et al. (1992). "Infant auditory temporal acuity: gap detection." Child Develoment **63**: 260-272.

White, L. (2014). "Communicative function and prosodic form in speech timing." Speech communication **63**: 38-54.

White, L., S. L. Mattys, et al. (2012). "Language categorization by adults is based on sensitivity to durational cues, not rhythm class." Journal of Memory and Language **66**(4): 665-679.

Whiteford, K. L. and A. J. Oxenham (2017). "Auditory deficits in amusia extend beyond poor pitch perception." Neuropsychologia **99**: 213-224.

Whiten, A., J. Goodall, et al. (2001). "Charting cultural variation in chimpanzees " Behaviour **138**(11): 1481-1516.

Wich, S. A. and H. de Vries (2006). "Male monkeys remember which group members have given alarm calls." Proceedings of the Royal Society of London B: Biological Sciences **273**(1587): 735-740.

Wich, S. A., M. Krützen, et al. (2012). "Call cultures in orang-utans?" PloS one **7**(5): e36180.

Wich, S. A., K. B. Swartz, et al. (2009). "A case of spontaneous acquisition of a human sound by an orangutan." Primates **50**(1): 56-64.

Wiener, N. (1948). "Cybernetics." Scientific American **179**(5): 14-19.

Wienicke, A., U. Häusler, et al. (2001). "Auditory frequency discrimination in the squirrel monkey." Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology **187**(3): 189-195.

Wier, C. C., W. Jesteadt, et al. (1977). "Frequency discrimination as a function of frequency and sensation level." The Journal of the Acoustical Society of America **61**(1): 178-184.

Wild, B., F. A. Rodden, et al. (2003). "Neural correlates of laughter and humour." Brain **126**(10): 2121-2138.

Will, U. (1997). "Two types of octave relationships in central Australian vocal music?" Musicology Australia **20**(1): 6-14.

Will, U. (2002). Oral memory in Australian Aboriginal song performance and the parry-kirk debate: a cognitive ethnomusicological perspective. Proceedings of the international study group on music archaeology, 10. E. Hickmann and R. Eichmann.

Williamon, A. and J. W. Davidson (2002). "Exploring co-performer communication." Musicae Scientiae **6**(1): 53-72.

Williams, G. (1966). Adaptation and natural selection: A critique of some current evolutionary thought. Princeton, Princeton University Press.

Williamson, V. J., F. Liu, et al. (2012). "Perception and action de-coupling in congenital amusia: Sensitivity to task demands." Neuropsychologia **50**(1): 172-180.

Williamson, V. J., C. McDonald, et al. (2010). "Faster decline of pitch memory over time in congenital amusia." Advances in Cognitive Psychology **6**: 15.

Williamson, V. J. and L. Stewart (2010). "Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem." Memory **18**(6): 657-669.

Wilson, D. and E. Sober (1994). "Re-introducing group selection to the human behavioural sciences." Behavioural and Brain Sciences **17**: 585-654.

Wilson, M. and P. F. Cook (2016). "Rhythmic entrainment: why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed." Psychonomic bulletin & review **23**(6): 1647-1659.

Wilson, M. and G. Knoblich (2005). "The case for motor involvement in perceiving conspecifics." Psychological bulletin **131**(3): 460.

Wilson, M. and T. P. Wilson (2005). "An oscillator model of the timing of turn-taking." Psychonomic bulletin & review **12**(6): 957-968.

Wilson, M. L., M. D. Hauser, et al. (2001). "Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees?" Animal Behaviour **61**(6): 1203-1216.

Wilson, M. L., M. D. Hauser, et al. (2007). "Chimpanzees (Pan troglodytes) modify grouping and vocal behaviour in response to location-specific risk." Behaviour **144**(12): 1621-1653.

Wilson, R. A. and F. C. Keil (2001). The MIT encyclopedia of the cognitive sciences, MIT press.

Wilson, S. and J. Pressing (1999). Neuropsychological assessment and the modeling of musical deficits. Music medicine and music therapy: Expanding horizons. R. Pratt and D. E. Grocke. Melbourne, University of Melbourne**:** 47-74.

Wilson, S. M., A. P. Saygin, et al. (2004). "Listening to speech activates motor areas involved in speech production." Nature neuroscience **7**(7): 701.

Wiltermuth, S. S. (2012). "Synchronous activity boosts compliance with requests to aggress." Journal of Experimental Social Psychology **48**(1): 453-456.

Wiltermuth, S. S. and C. Heath (2009). "Synchrony and cooperation." Psychological science **20**(1): 1-5.

Wimmer, H. and J. Perner (1983). "Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception." Cognition **13**(1): 103-128.

Wing, A. and A. Kristofferson (1973). "The timing of interresponse intervals." Perception and Psychophysics **13**(3): 455-460.

Wing, A. M. (1980). "Timing of movement phases of a repeated response." Journal of Motor behavior **12**(2): 113-124.

Wing, A. M. (2002). "Voluntary timing and brain function: an information processing approach." Brain and cognition **48**(1): 7-30.

Wing, A. M. and P. J. Beek (2002). "Movement timing: a tutorial." Common mechanisms in perception and action **19**: 202-226.

Wing, A. M. and A. Kristofferson (1973). "The timing of interresponse intervals." Attention, Perception, & Psychophysics **13**(3): 455-460.

Winkler, I., G. P. Háden, et al. (2009). "Newborn infants detect the beat in music." Proceedings of the National Academy of Sciences **106**(7): 2468-2471.

Winter, P., P. Handley, et al. (1973). "Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation." Behaviour **47**(3): 230-239.

Winter, P., D. Ploog, et al. (1966). "Vocal repertoire of the squirrel monkey (Saimiri sciureus), its analysis and significance." Experimental brain research **1**(4): 359-384.

Witt, S. T., A. R. Laird, et al. (2008). "Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis." Neuroimage **42**(1): 343-356.

Wittgenstein, L. (2010). Philosophical investigations, John Wiley & Sons.

Wittmann, M. and E. Pöppel (1999). "Temporal mechanisms of the brain as fundamentals of communication—with special reference to music perception and performance." Musicae Scientiae **3**(1_suppl): 13-28.

Wolff, P. H. (1967). "The role of biological rhythms in early psychological development." Bulletin of the Menninger Clinic **31**(4): 197.

Wolpert, D. M., K. Doya, et al. (2003). "A unifying computational framework for motor control and social interaction." Philosophical Transactions of the Royal Society of London B: Biological Sciences **358**(1431): 593-602.

Wong, P. C., A. H. Chan, et al. (2011). "The bimusical brain is not two monomusical brains in one: evidence from musical affective processing." Journal of Cognitive Neuroscience **23**(12): 4082-4093.

Wong, P. C., A. K. Roy, et al. (2009). "Bimusicalism: The implicit dual enculturation of cognitive and affective systems." Music Perception: An Interdisciplinary Journal **27**(2): 81-88.

Wood, B. (1997). "Ecce Homo - behold mankind." Nature **390**: 120-121.

Wood, G. (1988). "Further field observations of the palm cockatoo Probosciger aterrimus in the Cape York Peninsula, Queensland." Corella **12**: 48-52.

Woodall, W. G. and J. K. Burgoon (1981). "The effects of nonverbal synchrony on message comprehension and persuasiveness." Journal of Nonverbal Behavior **5**(4): 207-223.

Woodrow, H. (1930). "The reproduction of temporal intervals." Journal of experimental psychology **13**(6): 473.

Woodward, A. L. (1998). "Infants selectively encode the goal object of an actor's reach." Cognition **69**(1): 1-34.

Woodward, A. L. (1999). "Infants' ability to distinguish between purposeful and non-purposeful behaviors." Infant behavior and development **22**(2): 145-160.

Woolhouse, M. and I. Cross (2010). "Using interval cycles to model Krumhansl's tonal hierarchies." Music Theory Spectrum **32**(1): 60-78.

Wrangham, R. (2009). Catching fire: How cooking made us human, Basic Books.

Wray, A. (1998). "Protolanguage as a holistic system for social interaction." Language & communication **18**(1): 47-67.

Wray, A. (2000). Holistic utterances in protolanguage: the link from primates to humans, Cambridge University Press.

Wright, A. A., J. J. Rivera, et al. (2000). "Music perception and octave generalization in rhesus monkeys." Journal of Experimental Psychology: General **129**(3): 291.

Wright, R. L., J. W. Bevins, et al. (2016). "Metronome cueing of walking reduces gait variability after a cerebellar stroke." Frontiers in neurology **7**: 84.

Wynn, T. and F. L. Coolidge (2007). "Did a small but significant enhancement in working memory capacity power the evolution of modern thinking." In Rethinking the human revolution: new behavioural and biological perspectives on the origins and dispersal of modern humans. P. Mellars, K. Boyle, O. Bar-Yosef, and C. Stringer (eds.): 79–90. Cambridge: McDonald Institute for Archaeological Research.

Xu, Y., C. R. Larson, et al. (2004). "Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences." The Journal of the Acoustical Society of America **116**(2): 1168-1178.

Xygalatas, D., U. Schjoedt, et al. (2013). "Autobiographical memory in a fire-walking ritual." Journal of Cognition and Culture **13**(1-2): 1-16.

Yabe, H., M. Tervaniemi, et al. (1997). "Temporal window of integration revealed by MMN to sound omission." Neuroreport **8**(8): 1971-1974.

Yabe, H., M. Tervaniemi, et al. (1998). "Temporal window of integration of auditory information in the human brain." Psychophysiology **35**(5): 615-619.

Yamada, M. and T. Tsumura (1998). "Do piano lessons improve basic temporal controllability of maintaining a uniform tempo?" Journal of the Acoustical Society of Japan (E) **19**(2): 121-131.

Yantis, S. and A. P. Hillstrom (1994). "Stimulus-driven attentional capture: evidence from equiluminant visual objects." Journal of experimental psychology: Human perception and performance **20**(1): 95.

Yaqub, B., G. Gascon, et al. (1998). "Pure word deafness (acquired verbal auditory agnosia) in an Arabic speaking patient." Brain **111**(2): 457-466.

Yates, C. M., T. Justus, et al. (2017). "Effects of musical training and culture on meter perception." Psychology of Music **45**(2): 231-245.

Yin, P., J. B. Fritz, et al. (2010). "Do ferrets perceive relative pitch?" The Journal of the Acoustical Society of America **127**(3): 1673-1680.

Yu, L. and M. Tomonaga (2015). "Interactional synchrony in chimpanzees: examination through a finger-tapping experiment." Scientific reports **5**: 10218.

Yu, L. and M. Tomonaga (2016). "Unidirectional adaptation in tempo in pairs of chimpanzees during simultaneous tapping movement: an examination under face-to-face setup." Primates **57**(2): 181-185.

Zahavi, A. (1975). "Mate selection: A selection of handicap." Journal of Theoretical Biology **53**: 205-214.

Zahavi, A. (1997). The handicap principle. Oxford, Oxford University Press.

Zarate, J. M. (2013). "The neural control of singing." Frontiers in human neuroscience **7**: 237.

Zarate, J. M., S. Wood, et al. (2010). "Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers." Neuropsychologia **48**(2): 607-618.

Zarate, J. M. and R. J. Zatorre (2008). "Experience-dependent neural substrates involved in vocal pitch regulation during singing." Neuroimage **40**(4): 1871-1887.

Zarco, W., H. Merchant, et al. (2009). "Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys." Journal of neurophysiology **102**(6): 3191-3202.

Zatorre, R. J. (1988). "Pitch perception of complex tones and human temporal-lobe function." The Journal of the Acoustical Society of America **84**(2): 566-572.

Zatorre, R. J., P. Belin, et al. (2002). "Structure and function of auditory cortex: music and speech." Trends in cognitive sciences **6**(1): 37-46.

Zatorre, R. J., J. L. Chen, et al. (2007). "When the brain plays music: auditory-motor interactions in music perception and production." Nature reviews. Neuroscience **8**(7): 547.

Zendel, B. R., M.-É. Lagrois, et al. (2015). "Attending to pitch information inhibits processing of pitch information: the curious case of amusia." Journal of Neuroscience **35**(9): 3815-3824.

Zendel, B. R., B. Ross, et al. (2011). "The effects of stimulus rate and tapping rate on tapping performance." Music Perception: An Interdisciplinary Journal **29**(1): 65-78.

Zentner, M. R. and J. Kagan (1998). "Infants' perception of consonance and dissonance in music." Infant Behavior and Development **21**(3): 483-492.

Zhou, P., F. Sui, et al. (2010). Music therapy on heart rate variability. Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on, IEEE.

Zhou, W.-X., D. Sornette, et al. (2005). "Discrete hierarchical organization of social group sizes." Proceedings of the Royal Society of London B: Biological Sciences **272**(1561): 439-444.

Zillman, D. and S. Gan (1997). Musical taste in adolescence. In The social psychology of music. D. Hargreaves and A. North (Eds.): 161-187. Oxford, Oxford University Press**.**

Zucker, L. (1946). "Psychological aspects of speech-melody." Journal of Social Psychology **23**: 73-138.