Pre-print submitted to **Comparative Cognition and Behavior Reviews**

The natural history of musical rhythm: functional and mechanistic theories on the evolution of human rhythm cognition and the relevance of rhythmic animal behaviors

Corresponding author:

Gabriela-Alina Sauciuc, Lund University, Department of Philosophy / Cognitive Science

Corresponding address: Lund University, Department of Philosophy, Box 192, 221 00 Lund.

Phone number: (+46) 70 64 866 86

e-mail address: gabriela-alina.sauciuc@lucs.lu.se

Co-authors

Tomas Persson, Lund University, Department of Philosophy / Cognitive Science

Corresponding address: Lund University, Department of Philosophy, Box 192, 221 00 Lund

e-mail address: tomas.persson@lucs.lu.se

Elainie Madsen, Lund University, Department of Philosophy / Cognitive Science

Corresponding address: Lund University, Department of Philosophy, Box 192, 221 00 Lund

e-mail address: elainie.madsen@lucs.lu.se

Keywords: rhythmic behaviors, rhythm cognition, music evolution, functional theories, mechanistic hypotheses

25 **Abstract**

26 There has recently been a growing interest to investigate rhythm cognition in nonhuman
27 animals as a way of tracking the evolutionary origins of human musicality - i.e., the ability to
28 perceive, enjoy and produce music. During the last two decades, there has been an explosion
29 of theoretical proposals aimed at explaining why and how humans have evolved into musical
30 beings, and the empirical comparative research has also gained momentum. In this paper, we
31 focus on the rhythmic component of musicality, and review functional and mechanistic
32 theoretical proposals on abilities regarded as prerequisites for perceiving and producing
33 rhythmic structures similar to those encountered in music. For each theoretical proposal we
34 also review supporting and contradictory empirical findings. To acknowledge that the
35 evolutionary study of musicality requires an interdisciplinary approach, our review strives to
36 cover perspectives and findings from as many disciplines as possible. We conclude with a
37 research agenda that highlights relevant, yet thus far neglected topics in the comparative and
38 evolutionary study of rhythm cognition. Specifically, we call for a widened research focus
39 that will include additional rhythmic abilities besides entrainment, additional channels of
40 perception and production besides the auditory and vocal ones, and a systematic focus on the
41 functional contexts in which rhythmic signals spontaneously occur. With this expanded focus,
42 and drawing from systematic observation and experimentation anchored in multiple
43 disciplines, animal research is bound to generate many important insights into the adaptive
44 pressures that forged the component abilities of human rhythm cognition and their
45 (socio)cognitive and (neuro)biological underpinnings.

1. Introduction

There has recently been a growing interest to investigate rhythm cognition in nonhuman animals (henceforth ‘animals’), as a way of tracking the evolutionary origins of human musicality - i.e., the ability to perceive, enjoy and produce music (e.g., Honing, 2019, and references therein). Currently, there is wide agreement that musicality is a biocultural ability that encompasses a set of cognitive and emotional traits that support it. Consistent with this view, a **compositional approach** to the evolution of musicality is presently advocated by several scholars (Fitch, 2006; Justus & Hutsler, 2005; McDermott & Hauser, 2005; Patel, 2006), whereby the presence and variability of component traits of human musicality ought to be systematically examined across human cultures and in non-human species. This approach offers real potential for mapping musicality-relevant traits that are shared with other species, thereby providing a unique insight into issues related to how, when and why musicality has evolved in our own species.

Since **rhythm** (together with the use of discrete pitches) is empirically established as an absolute universal in human music (Savage et al., 2015), abilities related to *rhythm cognition* constitute an obvious candidate topic for the comparative study of musicality. Musical rhythm is typically defined as the systematic patterning of sounds in time, with emphasis placed on the regular (i.e., periodic) recurrence of sounds or sound groupings (as opposed to, e.g., unpatterned sound repetition). While definitions such as this emphasize the auditory modality, musical rhythm must also encompass the regular patterning of body movements in time, since, across cultures of the world, dance is universally coupled to music (Cross, 2001; Savage et al., 2015; Richter & Ostovar, 2015). Ethnographic records show, moreover, that music and dance as separate phenomena and separate concepts have emerged only recently in the cultural history of humans. Indeed, it is not unusual that so-called traditional cultures lack specific concepts (and words) for music and / or dance (Cross, 2001;

Fitch, 2006; Merker 2000, 2002; Nettl, 2015). Instead, musical behaviors (e.g., dance, song, tool-based sound making) occur as an indissoluble part of rituals or religious ceremonies, and are often named after the specific ritual or purpose they serve (Cross, 2001; Nettl, 2015). While this may contrast with how we customarily consume, experience and conceive of music and dance in post-industrial, contemporary societies, it is likely to be highly informative about how and why musicality evolved, as hunter-gatherer societies are often considered a more faithful reflection of the socio-ecological niche in which our ancestors evolved (Mithen, 2005). It is also revealing that evolutionary theories focused on the putative adaptive functions of *musicality* are often theories of *dance* evolution as well, insofar as their focus lies on the benefits of rhythmically moving together (as detailed further down, in sections 2 and 3).

Proposals that rhythmicity constitutes the most ancestral component of musicality date back to the origins of the debate on the evolution of musicality (e.g., Darwin, 1871), with suggestions that the impulse to move rhythmically results from a general appetite for exercise, or a need to synchronize during work chores, such as pulling nets or pounding grain (as reviewed by Dissanayake, 2009a). In contemporary theories, the view that musical rhythm emerged prior to melody is backed by archaeological evidence suggesting that the pitch- and breathing control needed for producing melodic sound sequences was possible at the earliest in *Homo heidelbergensis* (i.e., about one million years ago), as this species exhibited suitable musculature for such capacities (Montagu, 2017; Morley, 2013). In contrast, the musculature of ancestral apes does not hinder the expression of bodily and percussive rhythms, thus suggesting that rhythmicity was likely the earliest manifestation of a musical ability.

Comparative research on rhythm as being the most ancestral component of musicality has begun to take off only during the last decade. Up until now, comparative research has primarily focused on *audiomotor entrainment*, which is the ability to accurately synchronize

bodily movements to an exogenous auditory rhythm. Emerging evidence suggests that individuals of several species can (be trained to) align motor responses to metronomic and even musical stimuli (Honing, 2019, and references therein). Although limited, this evidence has rekindled the century-long interest in the origins of human musicality (Darwin, 1871), spurring an effervescence of speculations and hypotheses related to the evolution of human rhythmic abilities.

Traditionally, the theoretical debate around the evolution of musicality has revolved much around whether musicality is an adaptation, i.e., whether it has evolved in response to selective pressures from the socio-ecological environment of our ancestors, and thus fulfilled important adaptive functions for the species (e.g., Fitch, 2006; Huron, 2012). Alternative proposals are that musicality is an exaptation (i.e., a trait originally lacking adaptive benefits, but later co-opted in the service of an adaptive function), a spandrel (i.e., a by-product of other adaptations), or a technology (e.g., Dissanayake 2000, 2009a; Fitch, 2006; Huron, 2001, 2012; Kotz et al. 2018; Mithen, 2005; Patel, 2006; Pinker, 1997). Several scholars have recently argued that such debates are counter-productive, since musicality is not a monolithic trait, but encompasses a conglomerate of abilities with distinctive adaptive benefits and evolutionary trajectories (e.g., Fitch, 2011; Kotz et al., 2018; Mithen, 2006; Trainor, 2018). In the evolution of complex traits such as musicality, it is generally the case that adaptation, exaptation and spandrel are intertwined and iterated (Andrews et al., 2002). The alternative approach that has gained momentum during the last decade is to separately investigate constituent components of musicality in order to formulate specific hypotheses about their putative adaptive functions and underlying mechanisms, which are amenable to empirical testing in both human and animal populations (e.g., Fitch, 2006; Trainor, 2018).

In the following sections of this paper, we review the main functional (section 2) and mechanistic (section 3) theoretical proposals on the evolution of musical rhythm and the

relevant empirical evidence that supports each of them. Our ambition in these two review sections is to cover as many explanatory angles as possible in order to establish links, and invite potential synergies between a diversity of traditional disciplines and relevant areas of research. The review is intended to provide a background for our own theoretical perspective and the research agenda ensuing from it, as outlined in section 4. Our research agenda stresses the need to conduct systematic research on rhythmic abilities in a variety of species, and emphasizes the value of comparisons between human and non-human rhythmic abilities.

Accounting for functions and mechanisms in separate sections may, sometimes, lead to redundancies or appear forced, considering that these are complementary levels of explanation. On the one hand, functional theories address questions about why a trait (e.g., a behavior) has evolved, thus attempting to account for the selective pressures that led to the emergence of the trait and the adaptive functions that it putatively fulfilled, given those specific selective pressures. On the other hand, mechanistic theories address questions about how the adaptive function(s) of a trait are achieved, thus accounting for the mechanisms (neuroanatomical, endocrine, genetic, cognitive, socio-cognitive or motivational) that implement it. By necessity, however, theorizing about the adaptive benefits of a trait is heavily informed by mechanistic evidence. As a consequence, mechanisms theorized to support putative functions of rhythmic features may be shortly mentioned in section 2, but treated more in depth in section 3.

2. Functional theories and derived hypotheses

In today's world, we spend a good portion of our time immersed in music, whether we relax, exercise, drive, shop, watch movies, play computer games, participate in demonstrations, military marches, etc. Research suggests that the rhythmic component of music exerts a plethora of emotional, aesthetic and social effects on humans, which makes it a

powerful regulatory tool for the individual. This also applies at group level, with benefits that seemingly range from cementing trust and cooperation to facilitating group cohesion, and communicating cultural identity (as reviewed in section 2.3). It is far from clear, however, to what extent these individual and social benefits reflect the adaptive significance that rhythm-based behaviors had for ancestral Hominins.

The recent wave of empirical studies on animal rhythmicity has focused primarily on mechanistic hypotheses regarding the evolution of musical rhythm and, by extension, the evolution of musicality. Functional hypotheses - i.e., hypotheses pertaining to the adaptive benefits of specific traits and the putative selection pressures that motivated their emergence - are rare in this recent literature from the field of animal cognition and bio-musicality, but are more common in other fields (e.g., evolutionary biology, evolutionary psychology, ethnomusicology, music psychology, and cultural anthropology), which invoke comparative data when building up their arguments. It is important to stress that the majority of these functional hypotheses are generically formulated, in the sense that the envisaged trait is musicality in general (as if it were a monolithic trait), rather than rhythmicity in particular. However, the adaptive benefits in focus are frequently related to rhythmic displays and/ or synchronization effects, which entails that rhythmicity is likely the relevant aspect. In this section, we review the most prominent functional theories on the evolution of musical rhythm, as well as the empirical evidence they have generated (for a quick overview of these, including key references, see Table 1).

Insert Table 1 about here

2.1. Musicality as a sexually selected trait

Darwin (1871) suggested that, by analogy with other species, music and dance may

have originated in the mating displays of our ancestors, and speculated that the strong emotional impact that music exerts on humans owes to these courtship origins. He argued that both musical notes and rhythms are vehicles by which primeval displays advertised physical or mental fitness to the opposite sex. This theory has been revived and elaborated by Miller (2000). As an argument for the sexual selection hypothesis, Miller draws a parallel between the high complexity and energy expenditure of mating displays in other species and the musical behaviors of tribal human societies, both of which seemingly involve the largest and most energy-craving muscles in the human body, and commonly feature repeated high stepping, stamping, and jumping. Darwin's initial proposal did not imply a sex bias in the production of protomusical behaviors. Miller, however, sides with the idea that musical behaviors have emerged as male-specific display, with arguments that, e.g., music production is more widespread in men than women, and purportedly peaks in men around 30 years of age, which is claimed by Miller to be the prime age for courtship.

According to Miller (2000), musical behaviors function both as honest fitness signals and aesthetic displays. As an honest fitness signal, for example, a well-sustained rhythm - whether enacted in song, drumming or movement - would advertise aerobic fitness, health, strength, stamina, and good motor coordination, as well as cognitive abilities, such as memory, learning, and the processing of complex sequences. All these skills, it is argued, are valuable for persistence hunting, which involves party coordination and prolonged tracking of prey, and is currently practiced by hunter-gatherer communities, and, thus, was likely also practiced by our ancestors.

As aesthetic displays, musical behaviors would be signal traits shaped by evolutionary processes that exploited species-specific sensory biases, thereby setting up sexual-selection pressures in favor of behaviors that appealed to such biases. In the field of animal communication, this process is captured by the **sensory exploitation** hypothesis (e.g., Ryan et

al., 1990), according to which the preference for certain signal traits may be the outcome of a sensory bias that already existed in receivers before the signalers evolved the traits to exploit it.

An incontestable merit of Miller's comprehensive proposal is that it sketches several empirically testable hypotheses, and suggests exploratory analyses aimed at assessing the theory that musicality has evolved as a sexually-selected trait. Several of these hypotheses concern the predictive power of musical skills with respect to reproductive success and the physical or mental aptitudes purportedly advertised by musical skill. Miller's suggestions for exploratory research revolve around the hypothesis that musical behaviors function as aesthetic display, which, among other things, predicts that rhythmicity prompts an optimal level of neural excitation in the receiver. These exploratory analyses focus on inventorying stimuli that are optimally exciting - and, thereby, attractive for our species - and assessing whether such optimality ranges extend across closely related species.

Currently, the empirical evidence to evaluate the theory that musicality has evolved as a sexually selected trait comes from three areas of investigation: *ethnographic data (i)*, *experimental studies with humans (ii)*, and *animal communication studies (iii)*. The *ethnographic evidence (i)* is ambiguous, with both consistent and divergent records. For example, Malinowski (1929) and Sachs (1962) provide several examples of musical behaviors used competitively, as male display. Moreover, a recent study has revealed that, across cultures of the world, singing and instrumental music are predominantly performed by males (Savage et al., 2015). Counter-examples, however, are also abundant in ethnographic records. It is relatively common in traditional societies that older men and women, who are well-beyond the optimal reproductive age, are regarded as the most skilled performers. There are also communities, such as the Moso (from the provinces of Yunnan and Sichuan in China), where the use of song in courtship rituals is an exclusive female prerogative (Namu &

Mathieu, 2004). Finally, musical behaviors are found in a broad range of contexts, and serve a variety of purposes, which is inconsistent with the idea of functional specialization.

Experimental studies with humans (ii) aimed at testing the hypothesis that musicality has evolved as a sexually-selected trait are currently scarce, and their results are inconsistent. In support of the sexual selection hypothesis, some studies suggest that musical skill may function as a putative indicator of mental or sexual fitness. For example, both women and men have been reported to prefer sexual partners who demonstrate some music abilities (Tifferet et al. 2012), and women (but not men) give higher ratings of facial attractiveness and dating desirability after listening to music compared to a silent control condition (Marin et al., 2017). Moreover, Madison et al. (2018) found that women rated prospective long-term partners higher with respect to intelligence, health and social status, when these partners were associated with high-quality musical performance. In another study, Mosing et al. (2015) found that musical discrimination skills correlated with general intelligence, although this correlation was accounted for by genetic heritability in only males and not in females. This finding is inconsistent with the hypothesis that musicality has evolved as a sexually selected trait driven by female assessment of male displays. Moreover, in this study, musical skills generally failed to predict measures of mating success, and no genetic correlations between musical aptitude and mating success were found (Mosing et al., 2015). While in these studies components of musicality related to e.g., pitch or rhythm were not tested separately, there are indications that rhythmic synchronization is mediated by sex hormones (Miani 2016 a,b, see also 2.2. for more details), and that men show superior rhythmic synchronization skills (Miani, 2016b).

Findings from *animal communication (iii)* studies are often invoked in theoretical proposals related to the hypothesis that musicality - in particular, musical rhythm - has evolved as a sexually selected trait. In animals, the main functional context in which rhythmic

behaviors - primarily rhythmic sound production - have been investigated is that of mating displays. For many species, it seems that rhythmic displays have an exclusively reproductive function (Greenfield. 2005), as such displays seem to be strictly regulated by sex hormones (for a review, see Schlinger and Brenowitz, 2017), and are often a male-specific trait (ten Cate & Spierings, 2019; Party et al., 2014). It is important to note that scholars who invoke animal data as potential evidence to back up the theory that musicality has evolved as a sexually selected trait (or any other evolutionary theory of musicality, for that matter) do not claim that animal signals are musical. Instead, features of animal signals that are reminiscent of human musical behaviors (e.g., rhythmic structure and communicative functions) are studied as potential *precursors* and behavioral *analogues* meant to inform on the selective pressures that could be conducive to rhythm cognition and rhythmic communication, as it is ultimately found in human musicality.

Rhythmic signals have been studied most extensively in arthropod, anuran, and avian species, whose mating calls - also designated *songs* - have been traditionally likened to human music (Fitch, 2006; Ravignani et al., 2019a). In mammalian species, studies of the presence of rhythm in courtship displays have only recently begun to take off, with evidence coming primarily from marine mammals and bats. For example, throughout the breeding season, males of certain pinniped species and baleen whales (e.g., the bearded seal (*Erignathus barbatus*), harbor seal (*Phoca vitulina*), walrus (*Odobenus rosmarus*), leopard seal (*Hydrurga leptonyx*), blue whale (*Balaenoptera musculus*), humpback whale (*Megaptera novaeangliae*), minke whale (*Balaena rostrata*), fin whale (*Balaenoptera physalus*)), emit intense bouts of rhythmic calls underwater (e.g., Handel et al., 2009, 2012; Handel & Mercado, 2016; Rogers, 2017; Scheider & Mercado, 2019). Similar rhythmically structured calls that coincide with the breeding season have been described in the greater sac-winged bat (*Saccopteryx bilineata*, Burchardt et al., 2019; Knörnschild et al., 2017).

Given the lack of studies that detail both rhythmic features and relevant contextual parameters, it is currently not possible to determine why rhythmic calling has been evolutionarily selected in the courtship context specifically. It has been hypothesized that for species with a low-density distribution (e.g., certain marine and parrot species), call rhythmicity functions as a failsafe redundancy that preserves call distinctiveness and enhances its recognition by distant receivers. Such rhythmicity has thus evolved to counteract the challenges of long-distance communication by reducing information loss during transmission (e.g., Heinsohn et al. 2017; Rogers, 2017). Conversely, in species with high population density, it has been hypothesized that rhythmic calling counters signal masking in noisy environments (e.g., Brumm & Slater, 2006; Serrano & Terhune, 2001). The predictability of rhythmic signals may also enhance auditory processing, in that attentional allocation would be optimized to fit the recurrence of relevant acoustic events, as proposed by the **dynamic attending theory** (e.g., Large & Jones, 1999, and as further detailed in section 3.1). This, in turn, would be consistent with the hypothesis that, in whale song, rhythmicity may accomplish a mnemonic function by facilitating the memorization of new elements - although this hypothesis remains untested (Scheider & Mercado, 2019).

A common assumption in theories of musicality as a sexually selected trait is that the specific *ways* in which rhythmic signals are executed function as honest fitness indicators, thus informing female receivers of male caller capacities. For example, in marine mammals that call underwater where breathing is not possible, the quality of rhythmic calling might advertise the caller's breath-holding capacity, which in turn is related to endurance and, thus, hunting ability. This hypothesis has been recently corroborated by an empirical study which revealed that, in leopard seals, rhythmic calling is more consistent in large-sized males compared to smaller ones, with call stability decreasing in the latter as the breeding season progresses (Rogers, 2017). Similar hypotheses have been advanced for the rhythmic

vocalizing of humpback whales on breeding grounds. Since in this species male songs are updated annually with novel elements that spread across all males in a region (Eriksen et al., 2005), it has been hypothesized that the temporal consistency and structural complexity of a male's rendition of an updated song serve as fitness indicators for attracting females (Chu & Harcourt, 1986; Medrano et al., 1996; Tyak, 1981). An alternative hypothesis is that whale song serves as an inter-male competitive display for establishing a dominance hierarchy, based on age and status information provided by the features of individual song execution (Darling & Béruré, 2001). Finally, a third hypothesis that relates whale song to reproductive function is that whale song functions as a long-range sonar, allowing males to locate females (Frazer & Mercado, 2000), or more broadly to locate other individuals, for purposes not only related to mating, but also to traveling or foraging (Mercado, 2018).

In a number of orthopterans (e.g., grasshoppers, locusts and crickets) and anurans (frogs), rhythmic courtship takes the form of communal displays, often involving a large number of individuals that signal either synchronously or in alternation. The seemingly cooperative appearance of such displays, while occurring in an obviously competitive context, has intrigued scholars for a long time, compelling them to draw parallels to rhythmic synchronization in humans, and thus to human musicality (for discussions, see Ravignani et al., 2014; Wilson & Cook, 2016, and references therein). It has to be stressed, however, that there is a key difference between such animal signals and musical rhythm, in that the former are simple and innate, while the latter can take complex forms that are learned. Several hypotheses have been formulated to account for the synchronous rhythmic displays of orthopterans and anurans, with speculations that such communal calling improves the quality of the broadcasted signal, either through the preservation of species-specific rhythms, thus facilitating species recognition (the **signal conservation** hypothesis: Greenfield & Schul, 2008), or by maximizing peak signal amplitude, thus enhancing signal broadcast range (the

beacon effect hypothesis: Hartbauer et al., 2014; Morris et al., 1978).

Cooperation-focused hypotheses, as the above, are, however, undermined by acoustic analyses which reveal that, in several insect species, synchronous calling is the by-product of a competition to produce the leading signal. Moreover, playback experiments also show that females prefer the male that sings first. This has been explained as a form of ‘precedence effect’, whereby leading signals suppress attention to the subsequent signals, or as indicative of a preference for the more energetically expensive signals (Zahavi, 1975). Indeed, in certain species, the probability of producing a leader chirp is associated with higher energetic expenditure, although it fails to correlate with other fitness indicators (for a review and discussion, see Hartbauer & Römer, 2016, and references therein). Followers, on the other hand, while being less conspicuous to potential mates, are also less conspicuous to predators (Lee et al., 2009), which suggests an interplay of sexual and natural selection in shaping the dynamics of synchronous rhythmic calls. While cooperation-focused hypotheses fail to explain the communal displays of orthopterans and anurans, they may be valid when it comes to explaining the evolution of musical rhythm in humans. For example, Merker and colleagues have proposed that the selection pressure to attract migrating females led to the emergence of synchronized rhythmic displays in a human ancestor during the late Miocene (Merker, 2000; Merker et al., 2009). Synchronized stomping and vocalizing would have provided a **beacon** benefit in this context - i.e., increased sound intensity and conspicuousness - thereby enabling the signal to reach farther than single or unsynchronized signals.

2.2. Music as the emotional sibling of speech

Suggestions that music - in particular, song - and speech have common evolutionary origins date far back in the history of philosophical thinking, although these accounts do not focus specifically on rhythm (e.g., Brown, 2000b, 2017; Darwin, 1871; Jespersen, 1922;

Mithen, 2005; Roederer, 1984; Rousseau, 1761). Originally, the debate revolved around whether music or speech had precedence over the other (e.g., Darwin, 1871; Rousseau, 1761; Spencer, 1857). In contemporary times, this aspect of the debate is illustrated by proposals that musicality may have evolved as a by-product of speech-related adaptations, e.g., vocal learning. Hypotheses generated by such proposals have been empirically tested on comparative data, but their results remain inconclusive (as will be discussed in section 3.1.).

An alternative account, that can be traced back to Darwin (1871) and has been significantly developed by Brown (2000b; 2017) postulates that a rudimentary communication system eventually diverged into a referential communicative system (speech) and an emotional communicative system, i.e., music (for additional perspectives on this account, see also Mithen, 2005). Most arguments for the common origins of musicality and speech are theoretical, being the result of examining data from various areas of speech and music research in order to find similarities between the two. Such comparisons suggest that speech and singing engage overlapping neural networks (e.g., Musso et al., 2015; Özdemir et al., 2006), rely on similar mechanisms of emotional modulation (Brown, 2017; Filippi, 2016) and emotional induction (Ma & Thompson, 2015), exhibit certain structural similarities (such as generativity and hierarchical organization: Fitch, 2006), and attractiveness ratings for speaking and singing correlate within the same individual (Valentova, 2019). As critics have pointed out, however, such similarities may be the result of confounding variables, such as common neuromotor systems connected to auditory input and vocal output (Zatorre & Baum, 2012). This is highly plausible considering that ‘common origins’ accounts focus primarily on song (i.e., *vocal* musical behavior), overlooking other musical behaviors such as dance or tool-assisted sound production.

Beyond similarities related to common input and output systems, the production of song and speech diverges with respect to breathing patterns (Leanderson et al., 1987), the use

of the vocal apparatus (Sundberg 2018) and underlying processing mechanisms (Zatorre & Baum, 2012). Generally, singing is more demanding than speech, as it requires finer vocal control (Zarate, 2013), higher capacities to regulate breathing and subglottal pressure (Sundberg, 2018), as well as higher energy expenditure, due to engaging all respiratory muscles (intercostal, abdominal and diaphragm), as opposed to speech, which typically implicates only the intercostal muscles (Leanderson et al., 1987; Åkerlund & Gramming, 1994). Clinical evidence further suggests divergences with respect to the neuro-cognitive processing of speech and music, as speech processing is intact in patients with amusia (i.e., tone-deafness), while music processing is not impaired in aphasia patients (Mithen, 2005; Peretz & Coltheart, 2003).

These differences notwithstanding, a recent iterated learning study (Ma et al., 2019) has brought intriguing evidence that a single communication system can diverge, across several generations, under the exclusive pressure of communicative function. Acoustic analyses conducted in this study revealed that, across transmission chains, nonsense vocalizations described as fulfilling the function of communicating an emotional state evolved into vocalizations whose pitch and intensity levels diverged from average vocalizations, and exhibited a wider range of speech variation, and a tendency towards slower rate and longer duration. Moreover, these vocalizations were rated as more music-like by naive participants from two different cultures. Conversely, nonsense vocalizations whose function was described as referential, were rated as more speech-like. These findings are consistent with the ‘common origins’ theories reviewed above, as they suggest that demands for functional specialization may trigger a cascade of phonatory and acoustic changes ultimately driving a communication system to split into distinct systems. Since animal vocalizations function as behavioral expressions of emotional states (Bachorowski & Owren, 2003), this functional specialization could have been prompted by increased demands for

referential communication in ancestral human species. It is also possible, however, that such demands have more protracted origins, which extend to the last common ancestor of hominids, given evidence of referential vocal communication in the other great ape species (e.g., Crockford et al., 2014; O'Bryan et al., 2018).

2.3. Musicality as a biocultural adaptation in the service of group-level functions

It has been long suggested that musicality has emerged under selection pressures related to human sociality, supporting group functions with regard to within- and/or between-group cooperation or inter-group competitiveness (e.g., Kogan, 1997; Roederer, 1984). Consistent with these theories, cross-cultural evidence suggests that the rhythmic component of musicality may be inherently social, considering that, across cultures of the world, there is a universal association between isochronous rhythm (where events occur regularly, at equal time intervals) and musical *group* performance (Savage et al., 2015). While social theories on the evolution of musicality agree that rhythmic isochrony is the crucial feature that facilitates the social functions of musicality, they diverge with respect to the putative primordial function(s) ascribed to proto-musical behaviors. Some theories are relatively unspecific, in that musicality is assigned a generalized group-level function by reinforcing social life (Brown, 2000a; Koelsch, 2014; Loersch & Arbuckle, 2013; Merker et al., 2009; Roederer, 1984). Other theories emphasize a particular adaptive function as primordial for ancestral musical-like behaviors, with the most prominent among these being **social bonding** (Dunbar 2004, 2012; Huron 2001), **coalition signaling** (Bryant, 2013; Hagen & Bryant, 2003; Hagen & Hammerstein, 2009; Mehr et al., 2020), and **cooperation** (Brown. 2000a; Freeman, 2001; Kogan, 1997; Merker et al., 2009). Since social theories rely on arguments pertaining to human-specific features of sociality, most evidence invoked in support of such theories comes from human studies. Animal behaviors are, nevertheless, referenced by such theories insofar

as they are related to adaptive pressures presumed to have spurred the emergence of protomusical behaviors.

According to the **social bonding** theory (e.g., Dunbar, 2004, 2012; Huron, 2001), musicality has evolved as a prelinguistic mechanism that supplanted a common bonding mechanism of primates - social grooming - as group-size in early *Homo* species increased considerably. Primates can spend up to 20% of their awake time grooming each other, a behavior which triggers the release of endogenous endorphins and oxytocin (Crockford et al., 2013; Keverne et al., 1989; Machin & Dunbar, 2011). These neuropeptides are known for their anxiolytic (i.e., calming) effects, and have, thus, been suggested to act as a reinforcement mechanism to catalyze social bonding through grooming. Given the increased group-size of early *Homo* species, this dyadic bonding mechanism likely became impractical, as it required prohibitive time resources (Lehman et al., 2007). As a consequence, it has been postulated that novel bonding mechanisms, such as laughter, proto-musical chorusing and speech, emerged to replace grooming (Dunbar, 2004). Since, in *Homo* species, social bonding has putatively shifted from tactile to vocal means, this would have exerted a pressure on the development and refinement of the neuroanatomical mechanisms involved in such vocal bonding behaviors, such as the motor control of vocal output and bodily time keeping (Dunbar 2004, 2012). Early arguments in favor of this theory highlighted, for example, the potential of music to simultaneously reach, involve, and thus connect a large number of people (Huron, 2001). Moreover, the ancestral neurochemical mechanisms underlying primate bonding through grooming - endorphins and oxytocin - appear to also be activated by rhythmic behaviors (see section 3.4 for more details).

Several theories highlight musicality as a mechanism that enables individuals to move together in time, thereby supporting human-specific forms of **social cooperation**, besides the affiliative effects mentioned above (Brown, 2000a; Freeman, 2000; Kirschner & Tomasello,

2009; Merker et al., 2009;). Kogan (1997), as well as Merker and colleagues (2009) argued, for example, that the feelings of bonding engendered by music and dance contribute to group solidarity, and promote altruism, thus increasing the effectiveness of collective actions, such as defending against a predator or attacking a rival clan. Moreover, the rhythmic synchronization involved in musical behaviors entails a kind of intentional cooperation (Brown 2000a) or a motivation to share and co-represent mental states and emotional experiences (Kirschener & Tomasello 2009a), thus enabling individuals to jointly attend to and coordinate actions towards common goals (see section 3.3 for more details on such purported socio-cognitive and motivational mechanisms of musicality).

The experimental evidence in support of the **social bonding** theory and the **social cooperation** theory is abundant, as musical behaviors, such as communal singing, chanting, dancing and drumming, have a wide range of prosocial effects (Anshel & Kipper, 1988; Pearce et al., 2017; Reddish et al., 2013, 2014; Tarr et al., 2016; Wiltermuth & Heath, 2009) that fail to materialize when only passively listening to music (e.g., Dunbar et al., 2012). Moreover, these effects appear to be specifically triggered by the rhythmic components of music, as the prosocial effects of rhythmic synchronization extend beyond the context of music-making. Indeed, interpersonal rhythmic synchronization (e.g., rocking in chairs, finger tapping) leads to prosocial feelings, such as liking, rapport, trust, affiliation, entitativity (i.e., ‘feeling as one’), cooperation and generosity (Hove & Rise, 2009; Valdesolo et al., 2010; Wiltermuth & Heath, 2009). These effects emerge early in human development, which is also suggestive of ancestral evolutionary roots (e.g., Justus & Hutsler, 2005; McDermott & Hauser, 2005; Trainor 2018). For example, 12-month old infants prefer a person who rocks in synchrony with them as opposed to an asynchronous one, and this preference does not extend to synchrony exhibited by non-social entities (Tungenc et al., 2015). Fourteen-month old infants are more helpful after being engaged in interpersonal synchrony (Cirelli et al., 2014),

an effect which is present also in the absence of music (Cirelli et al., 2017). In preschoolers, the prosocial effects of interpersonal rhythmic synchronization are documented across a variety of behaviors, such as swinging, music-making, clapping and tapping (Kirschner & Tomasello, 2010; Rabinowich & Meltzoff, 2017; Tungenc & Cohen, 2016). Overall, these results show that, from an early stage of human development, rhythmic synchronization triggers a wide array of prosocial effects in both musical and non-musical contexts, thereby suggesting a social bonding function for the rhythmic component of musicality. In turn, this supports the componential view of the evolution of musicality. Since the early ontogenetic expression of an ability typically is taken as an indication of ancestral phylogeny, these results also support the view that rhythm is likely the oldest component of human musicality.

In the **coalition signaling** theory on the evolution of musicality, the relationship between interpersonal rhythmic synchronization and social bonding is considered to have a reversed causality, as compared to the **social bonding** theory. Rather than *facilitating* social bonding, proto-musical behaviors purportedly *advertised* the level of bonding within a group, with bonding being acquired by other means (Bryant, 2013; Hagen & Bryant, 2003; Hagen & Hammerstein, 2009; Mehr et al. 2020; Merker, 2000). A putative benefit of synchronous rhythmic displays is to confuse and frighten the enemy through the mimicry of a large animal (Merker, 2000). According to an alternative proposal (Hagen & Bryant, 2003), synchronous displays co-evolved with the complexity of social organization as a way of advertising group fitness in the formation of inter-group alliances. As such, complex and well-synchronized group displays purportedly signal internal stability and higher ability to act together and coordinate complex activities, since a high degree of synchronization entails long-term commitment and dedication.

Compared to the **social bonding** theory, the empirical evidence supporting the **coalition signaling** theory is more limited, and consists of findings suggesting that tight

synchronization may communicate high group quality (Hagen & Bryant 2003). More specifically, participants rated highly synchronized musical performances as being of better quality, and rated better performances as suggestive of higher coalition quality. Similar correlations have also been found beyond the context of music, as the formidability (size and muscularity) of a group was judged to be higher for audio-tracks of highly synchronized footsteps compared to audio-tracks of less synchronized footsteps (Fessler & Holbrook, 2016). Several experiments have also revealed that entitativity (the extent to which individuals are perceived as a cohesive group) is judged to be higher when individuals move in synchrony (Edelman & Harring 2014; Lakens, 2010; Lakens & Stel, 2011).

According to the **coalition signaling** theory, the crucial socio-ecological pressures that triggered the emergence of proto-musical behaviors are linked to two ancestral hominin traits: primate territoriality and social carnivory. Bryant, Hagen and colleagues have thus suggested that the vocal chorusing of certain social carnivores (e.g., lions, hyenas) and primates (e.g., gelada monkeys, chimpanzees) could be regarded as a potential precursor for this putative coalition signaling function of musicality (Bryant, 2013; Hagen & Bryant, 2003; Hagen & Hammerstein, 2009; Mehr et al., 2020).

Since coordinated calling with a territorial defense function has been also documented in monogamous primates (e.g., Caselli et al., 2014; Clink et al., 2020; Müller & Anzenberger, 2002) and non-migratory songbirds (Logue & Hall, 2014), another hypothesis has linked communal calling to ecological conditions that favor sedentary lifestyles and social stability (e.g., Logue & Hall, 2014). However, this is likely to be less relevant to evolutionary theories of rhythm cognition (and musicality), since the primary form of territorial calling in these primate and avian species is duetting, whereby the male and the female alternate their calls with high temporal precision, but without generating a periodic communal rhythm. The timing of such duet contributions is likely to be cued by local features announcing the end of

a sequence (as it is common for turn-taking in human verbal interactions) rather than an overall relational structure based on periodicity (for more details, see 3.2). Possible exceptions could be tarsiers and titi monkeys, where territorial calling can take the form of simultaneous chorusing that involves additional group members besides the breeding pair (Caselli et al., 2014; Clink et al., 2020).

While the chorusing of non-human species is speculated to rely on reflexive emotional contagion (as detailed in section 3.3), **a transition to (non-musical) intentional synchronization** allegedly enabled the emergence of synchronized proto-musical behaviors in the *Homo* lineage (Brown, 2007; Bryant, 2013). It is currently unclear if the chorusing of social carnivores and primates is subject to such intentional control - and if so, to what extent - but research on chimpanzee pant-hoot chorusing provides some intriguing findings in this respect. For example, chimpanzees flexibly adjust the duration of pant hoot elements in a way that favors chorusing (Fedurek et al. 2013; Mitani & Gros-Luis, 1998). Moreover, chimpanzees in neighboring communities appear to actively modify their pant-hoots to create community-specific calls (Crockford et al., 2004). Similar tempo adjustments have been documented in the coordinated calls that Indo-Pacific bottlenose dolphins emit when herding or defending mates (Moore et al., 2020).

2.4. Mother-infant interaction as the primordial context of proto-musical behaviors

The theoretical proposals that mother-infant interaction is the primordial context of musicality focus primarily on vocal behavior, and emphasize two putative adaptive functions that may explain why hominin vocalizations acquired music-like features: **mother-infant bonding** (Dissanayake, 2000, 2009a; Mithen, 2005) and **credible signaling of parental attention** (Mehr & Krasnow, 2017; Mehr et al., 2020).

According to the **mother-infant bonding** theory (e.g., Dissanayake 2000, 2009a),

music-like vocalizing originates in affiliative signals that were modified for the purpose of interacting with infants through a process of ritualization. The notion of ritualization is borrowed from theories on the evolution of animal acoustic communication, where it refers to a process by which instrumental behaviors evolve into communicative signals. In this process, they are gradually modified toward increased formalization, repetition, and exaggeration, so that they become prominent, distinctive and unambiguous (Grammer & Eibl-Eibesfeldt, 1990; Watanabe & Smuts, 1999). As such, it is argued that adult affiliative signals (e.g., head bobbing, body leaning toward, smiling, eye-brow flash, high-pitch voice modulations, etc.) were gradually modified into repetitive and exaggerated signals believed to be emotionally-evocative to infants by appealing to universal psychobiological sensitivities. This ritualised form of parent-infant communication would, thus, represent the primordial form of musicality, which was subsequently adopted in ceremonial contexts, where it was further elaborated through ritualization (Dissanayake, 2000, 2009a).

The **credible signaling of parental attention** theory proposes that infant-directed song, which is universally present across human cultures (Mehr et al., 2019; Trehub, 1993), was a primordial musical form (Mehr & Krasnow, 2017; Mehr et al., 2020). According to this theory, the origins of infant-directed song can be traced back to the contact calls of ancestral hominin species, which - besides identity and distance cues that are common in many species - have evolved to also encode parental cues of attention to the offspring. Purportedly, such attention cues, which caregivers would give when outside the visual field of the offspring, gradually acquired acoustic features that ultimately resulted in what we currently define as infant-directed song. The development of increasingly elaborated contact calls in the hominin lineage was purportedly driven by selection pressures that acted towards increased demands for parental investment (e.g., attention). On one hand, humans have a longer postnatal period of brain development and shorter interbirth intervals compared to the other apes, which results

in multiple immature offspring of various ages that likely simultaneously compete for parental attention (Mehr et al., 2020). On the other hand, bipedalism and loss of body hair in the human lineage prevent infants from clinging onto the parent's body, as the infants of other ape species do, thereby preventing constant proximity between mother and offspring. These shifts in parent-infant ecology purportedly prompted early *Homo* species to evolve song-like vocalizations that enabled foraging caregivers to continuously convey parental attention to their multiple offspring from some distance, without incurring foraging costs (Mehr & Krasnow, 2017).

It is important to note that the evolutionary scenario described above is very similar to a scenario that was proposed more than a decade earlier to account for the origins of speech (Falk, 2004). More specifically, speech was theorized to have emerged from infant-directed signals, which had evolved special features that enabled foraging parents to communicate with physically distant infants. Thus, the theory of speech evolution proposed by Falk (2004) and the theory of music evolution proposed by Mehr & Krasnow (2017) emphasize identical evolutionary pressures related to increased demands for parental attention, and identical adaptive solutions to deal with those demands, i.e., the elaboration of ancestral contact calls. The two theories, however, diverge with respect to the outcome of this evolutionary process: infant-directed speech versus infant-directed song. It is currently unclear how the two theories can be reconciled, since neither of them discusses when and what prompted the emergence of the specific features that differentiate infant-directed song from infant-directed speech. As current research shows, these two parent-infant communication systems are, indeed, acoustically distinct, and differ with respect to many pitch, timbre, and rhythmic features (Moser et al., 2020). Paralleling the 'musilanguage' theories reviewed in 2.2., one way to reconcile the two theories could be to regard infant-directed song and infant-directed speech as both resulting from an originally undifferentiated system of contact calls used to signal

parental attention to distally located offspring. This would, however, require that each of the two theories is complemented with additional assumptions concerning a putative differentiation stage and the adaptive demands that prompted it. In line with the theories reviewed in 2.2, other possibilities would be that one of the systems evolved from the other, or independently from one another but at different points in time. Both phenomena would then have survived because both have added to the survival of the species, or at least were attached to something that has.

Before concluding this section, it is important to note that the proposals reviewed in this section are primarily concerned with the evolutionary emergence of musical pitch and melodic modulation. The relevance of such proposals to the evolution of musical rhythm is, nevertheless, implied in the process of ritualization - i.e., the primary mechanism invoked to account for the emergence of infant-directed song. The explanatory power of evolutionary proposals focused on mother-infant interaction as the putative primordial context of musicality might benefit from extending their focus beyond vocal behavior. For example, for the purposes of infant soothing, caregivers from around the world prefer to use rocking rather than singing (Richter & Ostovar, 2016).

2.5. Therapeutic and ritual uses of musicality and rhythmicity

Self-report studies show that music is one of the most common affect-regulation strategies that humans use in everyday life to boost energy levels, or reduce tension (Saarikallio, 2011; Saarikallio & Erkkilä, 2007; Thayer et al. 1994). Music is also used in clinical settings, to alleviate stress-related disorders (Aalbers et al., 2017; Koelsch, 2009; Landis-Shack, 2017; Sihvonen et al., 2017; Thaut et al. 2015). Many studies show that the beneficial psychological effects of music are reflected in physiological measures, including heart rate, blood pressure and hormonal levels (for a recent review and meta-analysis, see de

Witte et al., 2020). Similar positive effects on stress levels, assessed with both physiological and behavioral measures, have been documented in multiple non-human species, including apes, monkeys, dogs, hens and rats (for a review, see Alworth & Buerkle, 2013), although the presence and type of such effects vary as a function of species and type of musical pieces.

It is currently unclear, however, which musical features - and thus mechanisms - underlie the clinical benefits of musicality, since relevant variables - such as tempo, rhythm, and musical content - are typically not dissociated in the design of intervention studies, or are inconsistently controlled in clinical and research practices (Aalbers et al., 2017; Leubner & Hinterberger, 2017; McPherson et al., 2019). Recent evidence, nonetheless, suggests that active interventions that implement rhythm-based activities (hand clapping, instrument playing, dancing) may have greater potential for improving physiological stress markers (McPherson et al., 2019). Additional findings show that greater stress relief is provided when listening to live music, compared to prerecorded music (for a review, see Witte et al., 2020).

The use of music as a coping strategy has been attributed to the sense of pleasure that arises when humans synchronize to external rhythms (Dunbar, 2012; Koelsch, 2014; Launay et al., 2016; Salimpoor et al., 2015; Thaut et al., 2015; Trost et al., 2017; Vuust & Kringelbach, 2010), which in turn may be linked to the anxiolytic and elating effects mediated by neurochemical mechanisms that implicate the release and circulation of dopamine, oxytocin, and endorphins (as detailed in section 3.4). Purportedly, these acted as reinforcers for the preservation and elaboration of rhythmic synchronization into rituals and, ultimately, into that which today is called music and dance (Brown, 2000a; Dissanayake, 2006, 2009b; Mithen, 2005). It is plausible that this reinforcement mechanism was already present in the last common ancestor of humans and chimpanzees, as suggested by the presence of entrainment in one type of chimpanzee behavior - the so-called 'Conga line' (Lameira et al., 2019). This behavior is a form of locomotion, whereby two or more individuals align in a row, either

646 standing or sitting, moving forward with a swaying movement, while maintaining bodily
 647 contact. Since Lameira et al. (2019) described the Conga-line as a stereotypical behavior, it
 648 has been speculated that rhythmic synchronization may have emerged in the *Homo* lineage as
 649 a coping co-stereotypy under the pressure of increasing socio-ecological stressors, such as
 650 resource scarcity (due to climate change) and crowding (as a consequence of increased group
 651 size and cooperative breeding). Additional data invoked by the authors to support this
 652 hypothesis includes data on chimpanzee rocking behaviors, which are commonly regarded as
 653 stereotypical behaviors related to confinement in captivity (e.g., Chamove, 1989; Lopresti-
 654 Goodman et al., 2013).

655 Other putative selective pressures that would have facilitated the recurrent use - and
 656 thus preservation - of musical behaviors as a coping mechanism may derive from enhanced
 657 cognitive abilities in the *Homo* lineage, in particular enhanced memory, foresight and causal
 658 reasoning abilities. Plausibly, this entailed an increased awareness of the uncertainty
 659 surrounding the outcome of future events, activities or natural phenomena, thereby leading to
 660 increased stress and anxiety. It has been proposed that apprehension about the future and the
 661 need to control uncertainty likely co-evolved with musicality - and the other artistic skills - as
 662 the motivators and the expression of what we currently define as religious practices (Brandt,
 663 2009; Dissanayake, 2009a,b). In other words, proto-musical behaviors - which took the form
 664 of ritualized versions of ordinary, instrumental behaviors - came to be connected to situations
 665 that had a vital significance to the group, as practices for curbing future uncertainty and
 666 mitigating collective crises. In such contexts, joining with others in rhythmic behaviors had
 667 purported coping benefits, by lowering stress and anxiety, and instilling a sense of
 668 uncertainty-control in the participants (Brandt, 2009; Dissanayake, 2009b; Mithen, 2005).

669 Consistent with the view that musicality - and in particular musical rhythm - has
 670 evolved to cope with future uncertainty, musical stimuli engage several components of the so-

called ‘default mode neural network’, which is commonly associated with remembering the past and simulating the future (Herdener et al., 2010; Toiviainen et al., 2020). In addition, musical stimuli improve memory functions and induce plasticity by increasing gray matter volume in a key structure of this network - the hippocampus (Herdener et al., 2010).

Strong support in favor of this view comes also from ethnographic records showing that musical behaviors are often linked to rituals that are associated with the supernatural, and have the function of invoking forces believed to affect vital aspects of group functioning, from hunting success and fertility, to protection from disease or warfare victory (Arom, 2000; Cross, 2003; Dissanayake, 2009b; Brandt, 2009; Nettl, 2015). As mentioned above, the widespread presence of synchronized displays in such ritualic practices has been attributed to the putative anxiolytic effects of rhythmic behaviors and, thus, to the sense of control that these may have evoked in situations in which ancestral humans had to deal with outcome uncertainty prior to crucial undertakings (Dissanayake, 2009b; Mithen, 2005). There are several examples of ritual practices that fit this scenario, such as the collective chanting and swaying of tribes in Papua New Guinea during violent storms (Malinowski, 1922; Mead, 1976(1930) apud Dissanayake 2009a), and the widespread use of communal musical behaviors in healing ceremonies and for lamenting the passing of group members (Dissanayake, 2009b; Mithen, 2005). Whether the anxiolytic use of collectively enacted rhythms constituted a primordial function in the *evolution* of musicality is, however, far from clear. In contemporary traditional societies, the ritual use of musical behaviors extends to a variety of rituals that are less clearly connected to a need for stress relief. As Seeger (1987:7) wrote about the lowland regions of South America: “wherever music is heard, something important is happening”. In perhaps all cultures of the world, musical behaviors are used - and conceptualized - as an integral part of important ceremonies, and as means to achieve important social purposes.

696

697 **3. Proximate mechanisms**

698 Many mechanistic hypotheses on the evolution of musicality focus on audiomotor
 699 entrainment, i.e., on the accurate synchronization of bodily movements to an auditory
 700 rhythmic stimulus. This is perhaps understandable, since audiomotor entrainment is regarded
 701 as the most iconic expression of human musicality, and has long been considered absent in
 702 other species. There is general agreement that entrainment relies on an ability to extract the
 703 temporal periodicity of an acoustic signal, and to predict its subsequent occurrence, which in
 704 turn enables individuals to synchronize their rhythmic movements with the acoustic signal.
 705 Mechanistic hypotheses on the evolution of other aspects of human rhythm cognition are
 706 generally lacking (as will be detailed in section 4).

707

708 **3.1 Neural adaptations underpinning rhythmicity**

709 One of the most productive hypotheses advanced within the new wave of comparative
 710 research on rhythmicity is the **vocal learning and rhythmic synchronization** hypothesis,
 711 henceforth **VLRS** (Patel, 2006). VLRS holds that audiomotor entrainment is a by-product of
 712 neural adaptations related to flexible vocal mimicry and, thus, to the evolution of human
 713 speech. In humans, this is supported by a series of similarities between music and speech with
 714 respect to rhythm processing and production. For example, the temporal organization of both
 715 language and music exploits perceptual grouping, whereby stimuli are mentally organized
 716 into hierarchically structured sets, whose boundaries are marked by changes in duration or
 717 frequency. Unlike speech, however, a quasi-universal feature of musical rhythm is its reliance
 718 on isochrony, i.e., on evenly spaced time intervals (Kotz et al., 2018; Ravignani & Madison,
 719 2017; Savage et al., 2015). In addition, music and speech differ in their mechanism of
 720 temporal coordination, with music being primarily a system of integration that requires

between-individual *coordination* of event *onsets*, as opposed to speech, which is a system of *alternation*, and requires the coordination of event *offsets* between individuals (Brown, 2007).

The **VLRS** predicts that only vocal learning species are able to synchronize to an auditory rhythm, while non-vocal learners will fail. Typically, songbirds, parrots, hummingbirds, seals, whales, bats and elephants are considered flexible vocal learners (for a review see Petkov & Jarvis, 2012), while other birds and mammals, including primates, are not (but see Lameira, 2017 for evidence of vocal learning in non-human apes). Initially, the comparative data appeared consistent with this divide, as eight budgerigars (*Melopsittacus undulatus*) could learn to synchronize motor responses with a metronomic stimulus (Hasegawa et al., 2011), and two parrots could synchronize head bobs with complex musical stimuli (Patel et al., 2009; Schachner et al., 2009). Moreover, a YouTube sampling study found that only individuals of vocal learning species (14 parrot species, one elephant species) exhibited audiomotor entrainment, while individuals of non-vocal learning species (including great apes) did not (Schachner et al., 2009). Initial experimental studies with rhesus macaques (*Macaca mulatta*) further suggested that non-human primates may be unable to synchronize motor output with an acoustic signal (Merchant & Honing, 2014).

The predictions of **VLRS** have, however, been contradicted by more recent evidence that non-vocal learners, such as rhesus macaques (Gámez et al., 2018; Takeya et al., 2017), sea lions (*Zalophus californianus*, Cook et al., 2013), and rats (*Rattus norvegicus domestica*, Noda et al., 2017) can learn to entrain. In addition, there is evidence of spontaneous entrainment in great apes (one chimpanzee, *Pan troglodytes*: Hattori et al., 2013; one bonobo, *Pan paniscus*: Large & Gray, 2015), whose status as flexible vocal learners is currently debated (for an extensive discussion see, e.g., Lameira, 2017). To account for such findings, an updated version of VLRS has been advanced, dubbed the ‘**intrinsic reward and rhythmic synchronization**’ hypothesis (Takeya et al., 2017). According to this hypothesis, vocal non-

learners may be able to acquire audiomotor entrainment, but only vocal learning species are motivated to display it *spontaneously*. The validity of **VLRS**, even in its updated form, is currently limited by the absence of evidence of entrainment in many vocal learning taxa (e.g., songbirds, whales, seals), and by evidence that cognitive mechanisms that enable entrainment in humans are absent in certain vocal learners, such as songbirds (for more details, see section 3.2 below). Moreover, with the exception of some parrot species, there is currently either no evidence of, or studies on, entrainment in other vocal learning species, including songbirds, whales and seals (Schachner et al., 2009).

Parallels between bird song and human song have a long tradition in theories on the evolution of human musicality, and, in fact, predate the VLRS hypothesis. For example, it has been pointed out that the vocalizations produced by birds during song are exclusively egressive (i.e., produced during expiration), which is similar to the great majority of human vocalizations. Moreover, similar descending melodic contours are present in both bird and human song (Savage et al. 2017). This has been further connected to anatomical and genetic similarities between the human and the avian vocal system (Pfenning et al., 2014; Savage et al., 2017). There are, however, several crucial differences with respect to anatomical structure, underlying physiology, neurocognitive processes and functional contexts that distinguish bird song from human song, thus limiting the usefulness of songbirds as a model species for understanding the evolutionary origins of human musicality. For example, unlike human musicality, which shows significant contextual flexibility, bird song is typically linked to either mating or territorial defense. Moreover, birdsong is highly dependent on steroid hormone release and hormone-neurotransmitter interactions, both during ontogenetic development and throughout the reproductive cycle. Songbirds appear to be, in fact, unique among vertebrates, in that they display considerable amounts of sex steroid receptors in non-limbic regions of the forebrain (Alward et al., 2017; Schlinger & Brenowitz, 2002). Such

receptors abound in neural assemblies that are implicated in song production, thus suggesting that sex steroids are directly involved in controlling birdsong. This high dependency on the release and fluctuations of sex steroids might explain the contextual inflexibility of bird song. This is not seen in other bird species, not even in other vocal learning birds, such as parrots or hummingbirds. Finally, neurocognitive mechanisms that are crucially implicated in the perception and production of musical rhythm are absent or, at best, only partially present in songbirds (for more details, see 3.2).

An alternative hypothesis, which, like VLRS, focuses on neural audiomotor adaptations (see Table 2 for an overview of hypotheses on neural adaptations that support rhythm production and perception), is **the gradual audiomotor evolution** hypothesis (in short, GAE). According to GAE, a progressive chain of anatomical and functional brain changes has led to the evolution of a human-specific beat-based timing (or relative timing) mechanism, in addition to the interval-based timing mechanism present in, e.g., monkeys (Merchant & Honing, 2014). As such, in the primate lineage, audiomotor entrainment is predicted to co-vary with the complexity of auditory and motor circuits and the tightness of their coupling. Monkeys are, thus, predicted to lack an ability for audiomotor entrainment, and to exhibit very few audiomotor neural connections. In contrast, nonhuman apes are predicted to exhibit higher complexity of relevant audiomotor circuitry, as well as rhythmic abilities that are superior to those of monkeys, but inferior to those of humans. Initial data appeared to support this hypothesis, as one chimpanzee (Hattori et al., 2013) and one bonobo (Large & Gray, 2015) were able to synchronize their motor responses (tapping, drumming) to auditory rhythms. Furthermore, this occurred spontaneously, in the absence of any training. Yet, the audiomotor alignment of these two apes was transient, and limited to stimuli that matched their preferred motor tempo of spontaneous tapping / drumming. Additional data suggests that musical rhythm triggers rhythmic movement in chimpanzees, a phenomenon

which is known as rhythmic engagement (Hattori & Tomonaga, 2019), and which, in human ontogeny, predates entrainment (Zentner & Eerola, 2010).

Insert Table 2 about here

The GAE is, however, challenged by several recent studies showing that entrainment abilities are not entirely lacking in rhesus macaques. As such, macaque monkeys can be trained to synchronize tapping movements with an isochronous metronome pulse (Gámez et al., 2018; Takeya et al., 2017), and can generalize isochronous tapping to non-trained tempi (Takeya et al., 2018). The monkeys are also able to adapt their tapping tempo to accelerating and decelerating metronomes (Gámez et al., 2018), and they modify the tempo of isochronous tapping by varying the inter-onset interval rather than the duration of the movement (Donnet et al., 2014).

Unlike the VLRS and GAE, other theories argue that entrainment is widespread in the animal kingdom. Cook et al. (2013), for example, propose that entrainment is enabled by **neural resonance** mechanisms that are largely conserved across taxa, and result from oscillatory interactions between neuronal populations in sensory and motor areas (e.g., Large & Gray, 2015; Rouse et al., 2016). During audiomotor entrainment oscillations in auditory neurons are believed to phase-lock on to the frequency of an exogenous auditory stimulus. In turn, this causes neural oscillations in the motor system also to phase-lock onto the same frequency, thereby leading to motor output that is entrained to the exogenous auditory input. Evidence from a study with rats suggests that, similarly to what has been observed in humans (e.g., Rouse et al., 2016), rhythm percepts based on neural resonance may indeed emerge in nonhuman brains (Noda et al., 2017), due to oscillatory interactions (such as those described above), but also due to additional neural mechanisms, such as **dynamic attending** (e.g., Large

& Jones, 1999) and **predictive coding** (e.g., Vuust et al., 2009).

The involvement of **dynamic attending** in rhythm processing is linked to the fact that, due to their regularity, rhythmic stimuli induce strong expectations about upcoming events. This causes attention to fluctuate in an oscillatory way, with most attentional resources being cyclically allocated to the most salient or most informative events in a sequence. Through dynamic attending, the deployment of highly energetic cognitive capacities is thus optimized to capture the most relevant stimuli. The process described above also entails that rhythmic stimuli are **coded predictively**, i.e., predictions are generated about event recurrence at regular points in time, which in turns enables optimal allocation of attention, as well as audiomotor entrainment (Vuust et al., 2009). Behaviorally, predictive coding is captured by the fact that bodily movements occur near the onset times of the auditory stimulus, with a slight anticipation of it (for reviews, see e.g., Fitch, 2013; Merchant & Honing, 2014; Merker et al., 2009; Repp, 2013). Since **dynamic attending** and **predictive coding** provide broad adaptive advantages, by optimizing stimulus reception and minimizing processing load (Large & Jones, 1999), and are crucially involved in rhythm perception and production, it is plausible that rhythmic signals are widespread in animal communication, as such signals would be easier to decode and recall, while also allowing for temporal pattern variations and, thus, species specificity.

Another proposal that argues for the widespread distribution of entrainment in the animal kingdom, and implicates neuroanatomical adaptations, is **the acoustic advantages** hypothesis (Larsson & Abbott, 2018). This hypothesis is an extension of an earlier version that specifically targeted the evolution of human entrainment as a consequence of **incidental sounds of bipedal locomotion** (Larsson, 2013; Larsson et al., 2019). Since bipedal locomotion entails periodic symmetric gaits, it purportedly enabled our ancestors to match each other's steps, and, thus, reduce footfall noise, which resulted in stealthier locomotion and

better detection of important environmental sounds. The **acoustic advantages hypothesis** extends this reasoning to vertebrates in general, proposing that the audiomotor circuits and genes involved in entrainment have their origins in fish schooling behavior and locomotor-respiratory coupling. In tetrapod descendants, this then evolved into couplings between forepaw motor processing and vocal communication, which favored synchronized movement and vocalization (Larsson & Abbott, 2018). Synchronized behaviors, such as schooling guided by incidental sounds of locomotion, putatively provided the adaptive advantage of improved situational awareness due to extended windows of silence. To explain evidence of entrainment in non-vocal learners, the **acoustic advantage hypothesis** contemplates the relationship between entrainment and vocal learning in reverse, i.e., the ability to entrain to external sounds is proposed to be a prerequisite for vocal learning (Larsson & Abbott, 2018).

A related hypothesis is the **bipedal experience in utero** hypothesis (Larsson et al., 2019; Parncutt & Chuckrow, 2017), whereby maternal walking provides coupled multisensorial (auditory, vestibular, tactile) and motor isochronous experience to the fetus, thereby facilitating the development of neural sensorimotor couplings. Hypothetically, this experience - and its underlying neural mechanism - would induce perceptual and affective biases for rhythmic stimuli in any of the three sensory modalities, with the effects being expected to be stronger when modalities are combined. Empirical research addressing evolutionary hypotheses of musicality connected to bipedalism is currently lacking. There are, however, some intriguing parallels between the optimal tempo of rhythmic synchronization (for a review see, e.g., Repp & Su, 2013) and the optimal walking tempo recorded in laboratory studies (MacDougall & Moore, 2005), as both are situated around 120 beats per minute (BPM). In addition, the tempo of popular dance music tends to cluster around 120-130 BPM (Leman et al., 2013). Humans, however, are able to synchronize with an auditory stimulus over a much broader tempo range (50 - 230 BPM), which exceeds the locomotor

(walking, running) tempo range, which is between 75–190 BPM (Larsson et al., 2019).

3.2. Cognitive mechanisms involved in rhythm perception and production

Currently, there is agreement that, through evolutionary time, musicality has been shaped by species-specific **bio-cognitive constraints** (e.g., Fitch, 2015; Honing & Ploegger, 2012; Huron, 2012; Merker et al., 2015). Candidate cognitive mechanisms discussed in the literature comprise both low- and high-level features, which range from **perceptual biases** (i.e., predilections for certain features) rooted in attentional, learning and memory processes to cognitive operations based on **relational processing**, such as hierarchical grouping, recursion and combinatorics (Brown, 2000b; Fitch, 2013, 2015; Merker et al., 2015; Ravignani et al., 2016; ten Cate & Spierings, 2019).

The involvement of some of these cognitive mechanisms in shaping musical rhythm is reflected by structural features that are universally found in music across cultures of the world, including (1) the presence of an isochronous beat; (2) the hierarchical organization of beats in a metric structure of strong and weak beats; (3) the predilection for hierarchical grouping based on (multiples of) two; (4) but also three beats; (5) the presence of motivic patterns (e.g., riffs) grounded on the beat (Savage et al., 2015). Interestingly, this set of universal features of musical rhythm has been replicated in the laboratory, in an iterated learning experiment, in which participants were allocated to several ‘generations’ in a transmission chain, and had to reproduce a sequence of percussive sounds for the next ‘generation’ of participants. At the end of the experiment, the random sequences seeded into the first ‘generation’ of participants evolved into rhythmic patterns that exhibited all the universal features of musical rhythm listed above, although they also exhibited ‘cultural’ (i.e., transmission chain) specificity (Ravignani et al., 2016). Using a similar approach, couched in a task where participants had to tap or vocalize seeded sequences, Jacoby & McDermott

(2017) found that both US and Amazonian participants favored rhythms with intervals characterized by durations based on small integer ratios (e.g., 1:1, 2:1), although the specific intervals exhibited by the two populations differed qualitatively.

Given the converging outcomes of cross-cultural and experimental research, it has been proposed that the rhythmic structure of music is shaped by mechanisms, processes and biases that generically constrain human cognition. For example, working memory constraints may be responsible for the emergence of features that arguably make stimuli easier to process and recall, such as isochrony, motifs and few durational categories (Ravignani et al., 2016). Thus, many of the cognitive traits that shape musical rhythm - and enable musicality - may have evolved for the purposes of **auditory scene analysis**, i.e., as mechanisms that enable parsing auditory input (Honing et al., 2015; Sperber, 1996; Trainor, 2018). The tendency to assign the basic beat to the lowest-pitched instrument in an ensemble illustrates this point, as it is likely rooted in inner ear adaptations that give low-pitch superiority for temporal processing (e.g., Hove et al., 2014; Nelson & Young, 2010). In contrast, it has been suggested that rhythmic isochrony, beat-based timing and entrainment are traits that cannot be explained as non-musical adaptations (Trainor, 2018).

The above argument (that rhythmic isochrony, beat-based timing and entrainment have evolved as musicality-specific traits) rests primarily on comparative data, which initially suggested that isochrony detection, beat-based timing, and entrainment is virtually absent in other species, or, when present, not based on spontaneous motivation. However, the evidence reviewed in 3.1., suggests that spontaneous entrainment and rhythmic engagement with musical stimuli *is* present in chimpanzee and parrot species. Moreover, isochrony detection has now been demonstrated in several species (e.g., rats, jackdaw crows, budgerigars), although the cognitive mechanisms that enable it in humans and other species may, to some extent, differ (ten Cate & Spierings, 2019).

In humans, isochrony detection has been postulated to rely on **relational processing**, i.e., the ability to extract and assess relations between relations (ten Cate & Spierings, 2019). In the case of isochrony detection, this entails that structural patterns in the auditory stream are detected based on relative - rather than interval - timing (e.g., Merchant & Honing, 2014; Teki et al., 2011; ten Cate & Spierings, 2019). In addition, the production and detection of metric structure arguably relies on even more complex forms of relational processing that entail **hierarchical grouping** and **recursion** (Fitch, 2013). Such relational processing abilities seem to enable individuals to extract the hierarchical structure of meter, which is based on recurring patterns of strongly and weakly accented beats, with strong beats being relatively louder or longer than weak beats. For example, a duple meter would be based on a recurrent *strong-weak strong-weak* pattern, while a triple meter would be based on a recurrent *strong-weak-weak strong-weak-weak* pattern.

Another cognitive mechanism suggested to enable meter perception is **perceptual grouping**, i.e., the propensity to spontaneously group stimuli into units based on, e.g., their proximity or similarity (Fitch, 2016; Patel, 2006). In humans, a common form of perceptual grouping in the auditory domain is the so-called ‘iambic-trochaic law’, whereby sound sequences with contrasting duration or intensity/pitch are spontaneously perceived as binary groups with two distinct patterns. When the contrast is durational, humans tend to perceive an iambic structure, i.e., a rhythmic pattern accented on the second sound. When the contrast regards pitch (high-low) or intensity (loud-soft), humans tend to perceive the reversed (i.e., trochaic) pattern, with emphasis on the first element (e.g., Hay & Diehl, 2007). These propensities appear to be innate in humans (Abboub et al., 2016), although the iambic bias is susceptible to influences from the linguistic background of the individual, while the trochaic bias is robust regardless of linguistic background (Iversen et al., 2008).

It is important to note that **relational processing** and **perceptual grouping** are not

music-specific abilities, but can occur in a variety of domains, including theory of mind, language, and visual processing. It is not yet clear if such abilities translate from one domain to the other, although, in humans, relational processing abilities, for example, have been found to correlate between the domains of language and theory of mind (Oesch & Dunbar, 2017). The comparative study of relational processing in the domain of rhythm is currently in its infancy. So far, research that specifically targets perceptual rhythmic abilities in nonhuman species have produced inconclusive results. Currently, there is suggestive evidence that starlings (Hulse et al., 1984), jackdaws (Reinert, 1965 cited by ten Cate & Spierings, 2019) and rats (Celma-Miralles & Toro, 2020a) may be able to detect isochrony based on the relative timing of auditory stimuli, i.e., on an ability to detect the regular recurrence of durational patterns at the global level of the whole auditory sequence. To some extent, this seems also to be the case for zebra finches and budgerigars, although these species seem to rely on a mix of both relational and local cues for isochrony detection (ten Cate et al., 2016). Laboratory rats seem able to recognize the rhythmic structure of a musical excerpt when its pitch structure is scrambled, although it is unclear if their performance is based on relational mapping (i.e., on global detection of a durational pattern) or on local cues (e.g., the duration of the first sound in a sequence: Celma-Miralles & Toro, 2020b). Evidence of perceptual grouping in the form of a trochaic (but not an iambic) bias has been found in rats (Mora et al., 2013) and zebra finches (Spierings et al., 2017). In addition, after exposure to relevant stimuli that varied in duration, rats displayed the iambic bias (Toro & Nespors, 2015). Based on this evidence, it has been proposed that the trochaic (but not the iambic) bias may be an evolutionarily ancient principle for sound grouping, and that the two forms of sound grouping (trochaic, iambic) may depend on distinct mechanisms (Spierings et al., 2017; Toro, 2016).

In addition to the research reviewed in the present section (3.2), the evidence of entrainment reviewed in 3.1. implies that animals that are capable of audiomotor entrainment

also perceive isochrony, since isochrony detection is assumed to be a prerequisite for synchronizing bodily responses to an auditory stimulus. Intriguingly, the only species that have so far shown spontaneous entrainment - chimpanzees and potentially parrots - are also among the few species that display an ability to map abstract relations. Indeed, evidence of relational processing in other species is currently limited to great apes (Hribar et al., 2011 and references therein), corvids (Smirnova et al., 2015) and parrots (Obozova et al., 2015), which can match abstract visual relations on the basis of their similarity. The great apes also exhibit some recursive abilities in the theory of mind domain, as they show an awareness of how others monitor their own attentional states (Hall et al., 2016) and a sensitivity to others' false beliefs (e.g., Krupenye et al., 2016). Finally, chimpanzees are capable of cross-modal audio-visual structural mapping, being able to map an abstract relationship from the visual to the auditory domain (Ravignani & Sonnweber, 2017), and can also represent distant dependencies between visual stimuli, e.g., by detecting the regularity of non-adjacent elements (Sonnweber et al., 2015).

Summing up, there is widespread agreement that both low- and high-level cognitive abilities are implicated in the perception and production of musical rhythm (see Table 3 for an overview of these and key references). Extant evidence suggests that some of these abilities support auditory scene analysis, and are shared with many vertebrates. Other relevant cognitive abilities, such as relational processing, have been attested only in, e.g., great apes, corvids and parrots, and are seemingly absent in species that are commonly regarded as musical (e.g., songbirds). Currently, it is unclear which cognitive mechanisms - if any - are specialized for musical rhythm production and perception, considering that all the mechanisms reviewed in this section seemingly serve general cognitive abilities that span several domains. Much work remains to be done in order to answer this question, including research aimed at clarifying the role of relational reasoning in enabling musical rhythm, as

well as research mapping the presence of rhythmic abilities in a number of relevant non-human species (as detailed in section 4).

Insert Table 3 about here

3.3. Socio-cognitive and motivational mechanisms

A recent version of VLRS dubbed the ‘**intrinsic reward and rhythmic synchronization**’ hypothesis, proposes that vocal learning abilities mediate the *motivation* to engage in rhythmic synchronization (Takeya et al., 2017). As such, in contrast to vocal non-learning species, which can learn to entrain after extensive training, certain vocal learning species would show an intrinsic motivation to engage in it. While this updated version of VLRS accommodates recent evidence that non-vocal learning species, such as rhesus macaques (Gámez et al., 2018; Takeya et al., 2017) and sea lions (Cook et al., 2013), can learn to entrain, it fails to explain the absence of spontaneous entrainment in songbirds (Schachner et al., 2009; ten Cate & Spierings, 2019). Thus far, the only species (besides humans) that have exhibited *spontaneous* audiomotor entrainment are chimpanzees (Hattori et al., 2013; Large & Gray, 2015), and potentially parrots (Patel et al., 2009; Schachner et al., 2009). While flexible vocal learning is not contested for parrots, it is currently debated for chimpanzees (as reviewed by Lameira, 2017). Intriguingly, however, a relevant - yet overlooked - trait that chimpanzees and parrots seem to have in common is socialization through rhythmic movement (for a more detailed discussion, see section 4.3).

Another motivational mechanism (see Table 4 for an overview of hypotheses concerning socio-cognitive and motivational mechanisms in the evolution of rhythm cognition) potentially implicated in rhythmic synchronization is **emotional contagion**, whereby affective cues exhibited by an individual trigger similar behaviors and corresponding

states in other individuals. For example, Fritz & Koelsch (2013) proposed that **acoustically mediated emotional contagion** is a homologous trait that, alongside song and drumming, has been conducive to the emergence of musicality in the *Homo* lineage. By enabling the synchronization of mood and motivations in large-sized groups of ancestral hominins, acoustically mediated emotional contagion has purportedly played a role in facilitating vital group functions (cooperation, communication, cohesion), thereby contributing to the development and establishment of rituals.

Insert Table 4 about here

Fritz & Koelsch (2013) discussed two types of vocalizations that elicit a contagious effect in nonhuman apes, and also exhibit acoustic features reminiscent of musical structure: the play pant and the chimpanzee pant hoot. The play-pant is homologous to human laughter, and, like laughter, facilitates social interaction (Davila-Ross et al., 2011; Matsusaka, 2004). While human laughter has contagious and prosocial effects at group level, great ape play-panting has, thus far, primarily been documented at a dyadic level. The pant hoot has been documented in various contexts, both affiliative and agonistic (e.g., discovering food, connecting with distant group members, patrolling, displaying). Typically, this vocalization spreads within the group, potentially amplifying arousal in a social - and indeed interactive - fashion. As argued by Fritz & Koelsch (2013), the outputs of musicality (i.e., music and dance) differ from such chimpanzee vocalizations in that they mediate a broader spectrum of emotional experiences (Juslin & Västfjäll, 2008; Koelsch, 2010; Koelsch et al., 2010), which are transmitted to groups far larger than the typical gatherings of nonhuman apes (Dunbar, 2012; Fritz & Koelsch, 2013).

With respect to relevant acoustic features, it should be noted that the play-pant and the

pant-hoot share the repetition of a motive, something that hypothetically increases the emotional contagiousness of such vocalizations. As reviewed above (2.1, 2.4 and 2.5), the presence of repetitiveness in signal structure is commonly attributed to ritualization, whereby an instrumental behavior is gradually modified through selection processes toward increased formalization and simplification, and thus toward increased communicative efficiency (Grammer & Eibl-Eibesfeldt, 1990; Watanabe & Smuts, 1999). Just like these vocalizations, traditional music also has a strong formal structure, which is generally driven by a few simple rhythms (Savage et al., 2015).

A similar mechanism (and ensuing effects) lies at the core of the **contagious heterophony** theory, which proposes that unsynchronized group calls (as displayed by a number of nonhuman social species) are a common precursor for musicality and speech (Brown, 2007). Unlike the **acoustically mediated emotional contagion** theory, the **contagious heterophony** theory sees the evolution of human musicality as an analogical rather than homological phenomenon. According to the **contagious heterophony** theory, a crucial point in this evolutionary process was the transition from reflexive vocal contagion (which is considered common in non-human species) to intentionally cooperative chorusing, which is characteristic to music and, thus, postulated as uniquely human. As pointed out in 2.3, however, some nonhuman species may exhibit intentionality during chorusing, as suggested by evidence of voluntary control over coordinated calls in chimpanzees and bottlenose dolphins.

Other putative motivational mechanisms theorized to be involved in human entrainment are the human **propensity for gathering to display** as a group (e.g., dance, sing: Merker et al., 2015) and **shared intentionality** (Kirschner & Tomasello, 2009). Owing to observations that chimpanzees in the wild exhibit excited communal displays, in which loud calls co-occur with drumming and exaggerated bodily movements (Geissman, 2000), Merker

and colleagues (e.g., Merker et al., 2015) argued that the motivation to gather and display in large groups likely constitutes a primitive trait of musicality, which was present in the last common ancestor of humans and chimpanzees. In the human lineage, however, such displays have arguably undergone a process of refinement and elaboration, whereby the pace of voice and bodily displays became increasingly regular, which in turn enabled rhythmic synchronization.

Shared intentionality, on the other hand, captures the purportedly human-specific motivation to share - and co-represent - mental states with others, which in turn enables individuals to jointly attend to and coordinate actions towards common goals. This kind of motivation is theorized to have emerged late in human evolution, as recently as 400 000 years ago, with the onset of obligate, collaborative foraging, which required increased and frequent joint coordinated action (e.g., Tomasello et al., 2012). The contribution of shared intentionality to the emergence of human musicality remains an understudied topic. Extant evidence - although very limited - suggests that synchronization-induced prosociality (as reviewed in 2.3) is the result of a co-dependence between shared intentionality and rhythmic synchronization, rather than being the effect of synchronization alone. On one hand, rhythmic synchronization produces prosocial effects only in contexts that prompt the representation of shared goals, i.e., when rhythmic synchronization is the result of a joint intention to synchronize, as opposed to when synchronization is the by-product of individuals separately synchronizing with the same beat (Reddish et al., 2013). On the other hand, across situations that impose shared goal representation, prosocial effects *only* emerge in the presence of rhythmic synchronization. For example, participants cooperated more when instructed to work together to synchronize with respect to phase and tempo, as opposed to when the participants were instructed to work together to keep the same tempo, but by delivering sequential, rather than simultaneous, contributions to rhythm keeping (Reddish et al., 2013).

Thus, it has been suggested that a **reinforcement of cooperation model** may explain the persistence of protomusical behaviors across evolutionary time. As such, perceived synchrony during the pursuit of shared goals would give an immediate feedback for achieving successful cooperation, thereby reinforcing the group's cooperative tendencies (*idem*).

Developmental evidence suggests that **shared intentionality** facilitates the expression of musicality - in particular rhythmic synchronization - in toddlers. Indeed, the age at which children exhibit flexible and accurate entrainment is lowered from 4,5 to 2,5 years of age if the entrainment task (e.g., drumming) is set in a social context (Kirschner & Tomasello, 2009). The presence of a social partner supposedly prompts a shared representation of the task as a joint action, thus engaging the motivation to co-represent the task and synchronize movements in time. Fundamental components of shared intentionality, such as joint attention and engagement in imitation games, emerge already at 6 months of age in humans (Reddy, 2015; Rossmanith et al., 2014; Sauciuc et al., 2020). Thus, it is also likely that shared intentionality mediates the prosocial effects of rhythmic synchronization documented in 12-, 14-, and 15-month old infants (as reviewed in 2.3), since such effects are specific to social stimuli. If the evolution of human musicality can be explained by its cooperative effects, the key mediating mechanism is likely shared intentionality. Shared intentionality has been theorized to be unique to humans, but recent evidence shows that all great ape species exhibit behaviors indicative of it. For example, in the context of social play, all great ape species engage in joint attention and use communicative gestures to re-engage reluctant partners, which in turn suggests an understanding of collaborative roles and shared motivations (chimpanzees, bonobos: MacLean & Hare, 2013; gorillas: Tanner & Byrne, 2010; orangutans: Gruber, 2013). Moreover, when being imitated, apes show signs of enjoyment and playfulness, as expressed by imitation games, play face, and laughter (Persson et al., 2018; Sauciuc et al. 2017).

1121

1122 **3.4. Neurohormonal mechanisms of musical behaviors**

1123 Research conducted within the last decade suggests that neurotransmitters, such as
 1124 dopamine, endorphin, oxytocin and vasopressin, are implicated in the production of musical
 1125 behaviors, as well as in mediating their positive effects. Several of the findings reported
 1126 recently relate specifically to rhythm perception or rhythmic synchronization, and may
 1127 account for mechanisms which, evolutionarily, promoted a predilection for rhythmic
 1128 communication and facilitated the cross-generational transmission - and thus preservation - of
 1129 musical behaviors (see 2.3, 2.5 for theories featuring such arguments). Indeed, all of these
 1130 neurotransmitters are known to promote psychological well-being through interactions with
 1131 the hypothalamic-pituitary-adrenal (HPA) axis (for reviews, see, e.g., Carson et al., 2013;
 1132 Pillozzi et al., 2020; Stanwood, 2019), which, in turn, are thought to mediate prosociality
 1133 during communal performance. These neurotransmitters are, thus, crucial for achieving the
 1134 adaptive benefits of musicality proposed by various functional theories on the evolution of
 1135 musicality (as reviewed in section 2).

1136 **Endorphins** are endogenous opioid neuropeptides produced in the pituitary gland,
 1137 which are released in response to pain, and have an analgesic effect. Endorphin release has
 1138 been linked to euphoric effects induced by physical exertion, and rewarding activities such as
 1139 sex, laughter, and eating (Chaudhry & Bhimji, 2018). The role of the opioid system in the
 1140 evolution of musicality - in particular the role of *endorphins* - has been emphasized by
 1141 proponents of the **social bonding** theory, based on evidence that rhythmically synchronized
 1142 behavior - during both musical or non-musical activities - may implicate this neurochemical
 1143 mechanism (Cohen et al., 2009; Dunbar et al., 2012; Tarr et al., 2015, 2016). Indeed, active
 1144 musical behaviors (such as communal singing, chanting, dancing and drumming) seemingly
 1145 trigger endorphin release (e.g., Dunbar et al., 2012; Tarr et al., 2015, 2016), while passive

listening to music does not (e.g., Dunbar et al., 2012). Rhythmic interpersonal synchronization outside of a musical context appears to have comparable effects, as, e.g., synchronized rowing produces higher levels of endorphins than solo rowing (Cohen et al., 2009).

The anxiolytic functions of rhythm have also been invoked by evolutionary theories that stress the therapeutic functions of musicality (as reviewed in 2.5), and have linked proto-musical motor behaviors to animal stereotypic movement (Dissanayake, 2009a; Lameira et al., 2019), which is known to result in reduced stress (Chamove, 1989). Brain imaging evidence suggests that music exerts a strong influence on several clusters within the amygdala that contribute to emotion regulation through the release of endorphins (Koelsch, 2014). While it remains currently unclear if the same neurochemical mechanisms mediate the anxiolytic effects of musical and stereotypic behaviors, it is plausible that the endorphin system is implicated in the *evolution* of proto-musical rhythmic displays in the primate lineage. Indeed, in the primate lineage the endorphin system is involved in regulating social bonding beyond kin and reproductive relationships (Chang et al., 2013; Launay et al., 2016), and spontaneous rhythmic synchronization with a potential anxiolytic function has been documented in chimpanzees (Lameira et al., 2019).

It has long been theorized that **dopamine** is involved in mediating the processing and production of musical rhythm, and the feeling of pleasure elicited by engaging with music (Ferreri et al., 2019; Salimpoor et al., 2015). Dopamine is generally known as the main neurotransmitter involved in reward and motivation processing, although distinct dopaminergic pathways are also implicated in learning, executive function, motor function and neuroendocrine control (Alcaro et al., 2007). Several cognitive computations that are dopamine-dependent have been proposed to account for the role of dopamine in musical behaviors, such as expectations regarding rhythmic structure and the violation of these

expectations (Salimpoor et al., 2015; Vuust & Kringelbach, 2010; Zatorre & Salimpoor, 2013), associative or episodic memory (Janata, 2009; Panksepp & Bernatzky, 2002), and temporal processing in the millisecond range (e.g., Merchant et al., 2013). Starting with the seminal study of Blood & Zatorre (2001), brain imaging research has repeatedly established increased activity in dopamine-rich areas, such as the striatum, when listening to pleasurable music (for reviews see Koelsch, 2014; Zatorre, 2015). Direct evidence for the role of dopamine in rhythmic motor control and synchronization (Koshimori et al., 2019; Miller et al., 2013) has only recently begun to emerge.

Freeman (2000) suggested **oxytocin** as an underlying mechanism of the social bonding function of musicality, while Feldman (2007) suggested that interactional synchronization triggers the release of oxytocin. Oxytocin - and the related neuropeptide **vasopressin** - are evolutionarily conserved neurochemical mechanisms that, in mammals, are generally implicated in aggression, as well as sexual and parental behaviors, through effects on subcortical structures in the forebrain and amygdala (for a review see, e.g., Rogers et al., 2018). In humans and chimpanzees, oxytocin and vasopressin have also been linked to social cognition and affiliation beyond sexual- or kin relationships (Crockford et al., 2013; Samuni et al., 2018; Webb et al., 2019), owing to effects exerted on cortical regions relevant to social reward, emotional recognition and empathy. Such effects have not been found in other primate species such as, e.g., macaques (Rogers et al., 2018).

Recent empirical research on the involvement of oxytocin and vasopressin in musical behaviors appears to confirm the early suggestions of Freeman (2000) and Feldman (2007), as singing in groups (Grape et al., 2002) and passive music listening (Nilsson, 2009) leads to an increase in peripheral oxytocin, especially for slow-paced rhythms (Ooishi et al., 2017). Moreover, therapeutic interventions that consisted of group drumming sessions were found to result in increased oxytocin levels and well-being in maltreated children (aged 8-12 years,

Yuhi et al., 2017). Conversely, intranasally administered oxytocin increases interpersonal rhythmic synchronization (Gebauer et al., 2014). This latter finding is potentially paralleled by a similar effect by **vasopressin**. Indeed, there are some indications that sexual arousal in males - which is mediated by increases in vasopressin - improves rhythmic synchronization (Miani, 2016a).

Insert Table 5 about here

As reviewed above, recent evidence suggests that several neurotransmitters are implicated in the production and/or perception of musical rhythm (see Table 5 for an overview and key references concerning the neurohormonal underpinnings of musical rhythm). Given the well-established role of these neurotransmitters in regulating a wide range of socio-behavioral traits, the evidence reviewed in this section seems to support social theories on the evolution of musical rhythm. Additional studies are, however, required to pinpoint the specific role that each of these neurotransmitters plays in the production and appreciation of musical rhythm. Indeed, in the formation and expression of socio-affiliative behaviors, oxytocin and vasopressin interact with dopamine, which is believed to exert a reinforcing influence on behaviors conducive to, e.g., affiliation (Baskerville & Douglas, 2010; Carson et al. 2013; Riedl & Javor, 2012). Similar interactions are likely to characterize the production and appreciation of musical rhythm, considering the bidirectional relationships that exist between rhythmic synchronization and social bonding (as reviewed in 2.3). Moreover, given the co-localization of the dopaminergic and opioid systems, it is likely that the hedonic effects of music-related behaviors are due to an interaction between the two systems, and, possibly, also with the oxytocinergic system.

3.5. The genetic basis of musical rhythm

The heritability of rhythmic abilities has been assessed in twin studies and with molecular genetic methods. The twin studies have yielded somewhat contrasting results. Ullén et al. (2014) found evidence of moderate (i.e., 50%) heritability for rhythm perception in a task that required participants to make similarity judgements for rhythmic strings. In contrast to these findings, Seesjärvi et al. (2015) found that rhythm perception, as measured by the ability to detect off-beat sounds, was best predicted by non-shared environment effects, as opposed to genetic or shared environment effects. It is possible that these inconsistent findings reflect distinct task requirements for measuring rhythm perception ability. As such, similarity judgments are likely to rely more on innate abilities for **auditory scene analysis** (as discussed in 3.2), and even on mechanisms that do not necessarily require **beat-based timing** (e.g., ten Cate & Spierings, 2019). In contrast, off-beat detection is likely to capitalize more on experience and acquired knowledge. Additional studies probing the heritability of the mechanisms that underpin the emotional processing of musical rhythm, in particular dopaminergic striatal functions and music-induced rewardness, suggest that these do not appear to be affected by genetic or shared environmental effects (Menne-Lothmann et al., 2012; Stokes et al., 2013).

Recent genome-wide association studies suggest that rhythmic skills display modest heritability. As such, the heritability estimates for the ability to synchronize to a musical rhythm range between 13-16% (Niarchou et al., 2019), while the heritability estimate for rhythmic discrimination skills was found to be 21% - with the latter result initially reported by a small-scale study, and subsequently replicated by a large-scale genomic wide-linkage assay study (Oikkonen et al., 2014; Pulli et al., 2008).

This pattern of findings is also reflected in findings from genomic analyses in humans, which highlight that genes implicated in the development of auditory pathways, sound

production, as well as non-specialized cognitive processes (such as learning, memory and attention) are important for musical skill (Mosing et al., 2014; Oikkonen et al., 2015). Such molecular mechanisms that are implicated in musical skill, but are not necessarily specific to musicality, appear to be ancient and well-conserved in vertebrates (Oikkonen et al., 2016). Genetic correlation analyses further indicate that rhythmic synchronization is a highly polygenic trait (i.e., a trait whose phenotype is influenced by many genes) that appears to share molecular architecture with traits relevant to emotion, cognition, neurodevelopment and health (Niarchou et al., 2019). As such, better rhythmic synchronization correlates with better lung function, faster processing speed in matching tasks, stronger hand grip, and the evening chronotype (i.e., people who prefer activity late in the day). Moreover, genomic loci related to neural structures such as cerebellum, basal ganglia and cortex contributed to the heritability of rhythmic synchronization, which is consistent with brain imaging findings on the neural underpinnings of rhythmic abilities. Finally, there are some indications that variations in the gene coding for the 1a receptor of the neuro-hormone arginine-vasopressin (which is widely implicated in regulating social behavior and socio-behavioral traits - as discussed in 3.4), is associated with traits related to musical rhythm, such as creative dance (Bachner et al., 2005), rhythm perception and rhythm memory (Granot et al., 2007; Ukkola et al., 2009).

Currently, very little is known about the genetic underpinnings of traits related to musical rhythmicity, with the handful of studies conducted so far targeting primarily rhythm discrimination (see Table 6 for an overview of current findings and key references concerning the genetic bases of musical rhythm). Future studies will thus need to investigate the heritability of additional rhythmic abilities, and uncover their molecular underpinnings. Since a small inventory of rhythmic features are universally present in musical traditions across the world, and since these seem connected to specific perceptual and production skills, it is highly plausible that genetic factors play an important role in the variability of their expression.

Genetic studies targeting each of these skills - as well as their interplay - will provide valuable data for research on the origins and evolution of human rhythmicity and musicality.

Insert Table 6 about here

4. A theoretical and empirical agenda for research on animal rhythmicity

As outlined above, the current status of research on the evolutionary origins of musicality, and musical rhythm, in particular, is characterized by an effervescence of theories and a mosaic of fragmentary evidence. Given this motley empirical background, researchers have launched repeated calls for systematic data collection, in which various animal taxa are emphasized as highly relevant for an evolutionary inquiry into the origins of musical rhythm and associated behaviors. Recurrent taxa in these calls are seals, whales, parrots and songbirds (Fitch, 2006; Kello et al., 2017; Kotz et al., 2018; ten Cate et al., 2016), thus implicitly suggesting underlying assumptions about convergent evolution. In other words, these calls imply that the rhythmic abilities involved in musicality may have repeatedly evolved in various species due to socio-environmental pressures similar to those that influenced their evolution in humans. Proposals favoring a homology-based perspective - i.e., that emphasizes the phylogenetic continuity of musicality-relevant traits in the primate lineage - are less common. Nevertheless, given the current compositional approach to the evolution of rhythmic abilities, phylogeny and convergence need not be mutually exclusive. For example, the phenomenon of tempo adaptation during communal calling in a cooperative context has been described in both chimpanzee (Crockford et al., 2004; Fedurek et al., 2013; Mitani & Gros-Luis, 1998) and dolphins (Moore et al., 2020). This could suggest that, in humans, tempo adaptation during communal sound-making is a phylogenetic trait ostensibly present in the last common ancestor of humans and chimpanzees. At the same time, the trait is also shared

with dolphins, a species in which it presumably evolved by analogy, given similar evolutionary constraints related to cooperation. However, considering the current empirical status of the field, it is premature to jump to any conclusions. In the next section, we will outline several research topics that need to be consolidated empirically - and in a systematic manner - if we are to make any major progress in disentangling the evolutionary roots of human musicality.

4.1. Musical rhythm is not a monolithic ability - moving beyond audiomotor entrainment

The empirical evidence reviewed in sections 2 and 3 suggests that synchronizing to a musical rhythm is not a monolithic ability. Likewise, musical rhythm is not a monolithic property of music. Instead, humans exhibit a suite of rhythmic abilities, which fall along a continuum of perceptual and motor sophistication. This point is well-illustrated by the protracted development of rhythmic abilities in human ontogeny, with perceptual rhythmic abilities generally preceding rhythmic production. With respect to perceptual abilities, experimental evidence suggests that neonates exhibit rudimentary beat perception (Winkler et al., 2009), i.e., ‘the ability to perceive a regular pulse in a temporally and/or acoustically varying rhythm’ (Honing, 2012). Beat perception was assessed by measuring neural oscillations while the neonates were exposed to drumming patterns that occasionally contained deviant structures, i.e., sequences in which certain sounds were omitted. The neonates appeared to have formed expectations about rhythmic structure, as they reacted primarily to structural changes that affected the strong beat (Winkler et al., 2009). Another milestone in the development of perceptual rhythmic abilities occurs at 3-5 months of age, when infants can learn to discriminate multi-modal (audio-visual) rhythmic patterns (Bahrick

& Lickliter, 2000; Bahrick et al., 2002). Around 7 months of age, infants begin to categorize rhythms based on their underlying beat (Hannon & Johnson, 2005), and at 8 months of age they begin to detect metric asynchrony in complex audio-visual stimuli, an ability which is significantly consolidated at 12 months (Hannon et al., 2017).

With respect to production abilities, infants begin to sparingly show rhythmic engagement (i.e., they display rhythmic movement when exposed to musical stimuli, and adapt bodily responses to the faster or slower tempo of the stimulus) from 3 months of age (e.g., Fuji et al., 2014). Accurate audiomotor entrainment, however, is only attained at 4 years of age, being generally restricted to stimuli that match the children's spontaneous tempo. As reviewed in 3.3, the age when this ability is expressed can be lowered to 2,5-years if the exogenous stimulus is embodied by a social partner. Adult-like proficiency with respect to timing accuracy and tempo flexibility is only attained at 8-10 years (as reviewed by Hannon et al., 2017), an age when children have also formed stable representations of the metrical structures present in their culture (e.g., Hannon et al., 2012).

In the light of this evidence, we argue that comparative and evolutionary research will benefit from considering the full set of quasi-universal features of musical rhythm, as well as the full spectrum of rhythmic abilities as they unfold in human ontogeny. As reviewed in 3.3, animal studies initially focused mostly on entrainment, using methods that required animal subjects to produce a motor response (tapping, head bobbing, etc.), and learn to synchronize it with an exogenous stimulus. Developmental data indicate that, in our species, accurate audiomotor entrainment is only achieved after years of practice and exposure to (auditory) rhythms. Thus, when probing rhythm cognition in nonhuman species, a heavy focus on motor responses that need to be acquired relatively quickly (i.e., within the timespan of an experiment), most likely underestimates the rhythmic abilities of those species, in particular those related to rhythm perception.

Recent studies of rhythm perception with a focus on isochrony detection are revealing in this respect. Initial findings suggest that isochrony detection - which is considered to be ‘the first cognitive step for beat perception’ (Celma-Miralles & Toro, 2020a) - may be a widespread trait, but it is possible that different mechanisms underlie this ability in different species (e.g., Celma-Miralles & Toro, 2020a; Honing et al., 2018; ten Cate & Spierings, 2019). For example, rhesus macaques, rats, zebra finches and budgerigars seem able to perceive isochrony across a variety of stimuli. The zebra finches, however, appear sensitive to stimulus duration and not the inter-stimulus interval, as it is characteristic for human musicality (ten Cate & Spierings, 2019). Rhesus macaques, on the other hand, detect the regularity of both visual and auditory stimuli (Gámez et al., 2018; Merchant et al., 2015; Zarco et al., 2009) by attending to inter-stimulus intervals, but they seem to do so by relying on absolute interval perception (Honing et al., 2018; Merchant & Honing, 2014) as opposed to human adults and infants who rely on beat perception (Bouwer et al., 2014; Winkler et al., 2009). Indeed, unlike human participants who react stronger to downbeat alterations, the monkeys seem to react equally to all alterations, regardless of their position (Honing et al., 2018). Finally, there are also results suggesting that rhesus monkeys can detect rhythmic groupings, as they show a differential - and stronger - reaction to omissions at the onset of a rhythmic group (Selezneva et al., 2013).

Observational studies and playback experiments can also be valuable for gaining insights into the kind of rhythmic patterns and structures that different species can discriminate. So far, the emphasis has been on the need for detailed analyses of multi-syllable, complex calls, that potentially exhibit a hierarchical structure, and that have been likened to music in the past, e.g., bird, whale or gibbon song. We maintain that the focus of such detailed analyses needs to be broadened to encompass vocalizations that we do not perceive as melodic, but which nevertheless exhibit a rhythmic structure, as well as non-vocal

rhythmic behaviors. Such studies are fundamental for understanding the perception and production capabilities of non-human species with respect to rhythmic structures of variable complexity, and for assessing if structural features of the temporal organization of animal signals bear similarities to musical rhythm. *Rhythm production capabilities* that are regularly involved in species-specific communication can be directly extracted from observational data with the help of appropriate analytic tools. *Perception capabilities* can sometimes be inferred from the data. For example, when the data contains sufficient signal variability with respect to given features (e.g., tempo, isochrony, timbre) and information concerning conspecific responses, it should be possible to determine which rhythmic features - if any - are likely to be salient to individuals of a given species, as they induce a measurable response. However, playback experiments, where individuals are presented with altered versions of a call, whereby single features are carefully and systematically varied, would be most useful in this respect.

Acoustic analyses of rhythmic features have recently been conducted on the calls of several marine mammal species (humpback whales: Schneider & Mercado, 2019; killer whales: Kello et al., 2017; Northern elephant seals: Mathevon et al., 2017; harbor seal: Ravignani et al., 2019b), birds (e.g., Kello et al., 2017; Roeske et al., 2020) and primates (e.g., de Gamba et al., 2019; Schruth et al., 2019; Terleph et al., 2017). Comparative studies of several species are also under way. For example, a comparison of the complex vocalizations of whale (humpback whale, killer whale) and bird (nightingale, hermit thrush) species with human speech and various music genres have revealed important differences (Kello et al., 2017). More specifically, this comparative study showed that hierarchical grouping and recursion were specific to human music, in which a small inventory of notes is combined in repetitive structures with a nested clustering of subunits. Bird and humpback whale song, on the other hand, relied on adjacency relationships, whereby only adjacent phrases shared sound

units, but distant ones did not. Killer whale vocalizations had features more similar to the structure of human conversation, in that they exhibited features of turn-taking coordination, which is comparable to findings from a study on duet coordination in gibbon song (Terleph et al., 2017).

Data from another study revealed, however, that the song of thrush nightingales exhibits a number of similarities to human musical rhythm at the lowest level of rhythmic organization, i.e., when comparing two successive inter-onset intervals in a sound sequence (Roeske et al., 2020). As such, both human music and thrush song showed comparable amounts of isochrony and a tendency towards rhythms that favor small integer ratios, although in thrushes this tendency was not significant. High-ratio rhythms, on the other hand, were very frequent in thrushes, while in human music high-ratios typically function only as ornaments. Finally, in both thrushes and human music the number of rhythm categories decreased with faster tempo, to ultimately transition into stereotyped isochrony. However, the threshold of this transition was much lower in thrush song (135 ms) compared to human music (265-410 ms), thus suggesting that only thrushes produced variable rhythms at tempi that equated the limit imposed by species-specific biomechanical constraints. In contrast, in human music, isochrony was present at tempi that were much slower than such constraints.

Worth of mention are also the findings of a comparative study that assessed the musical quality of the calls in 45 primate species (Schruth et al., 2019). Although rhythm was one of six musical variables scored in this work, this variable was subsequently incorporated into a composite index of “acoustic reappearance diversity”, which the authors used as a proxy for musical quality. This index was intended to capture the balance of redundancy (internal repetition) and diversity considered to be characteristic of human music (Brown & Jordania, 2013; Nettl 2015), by quantifying the number of distinct syllables in a call, as well as their reappearance. When pitted against environmental and social variables, this composite

index was found to be higher for species with an arboreal lifestyle, as well as for species that live in monogamy or small groups (2-6 individuals). These results are intriguing, as they run against several of the theories presented in section 2 and 3. Most notably, the **social bonding** theory (section 2.3) attributes the emergence of musicality to selective pressures imposed by increased group size (50+) in the human lineage. Other theoretical proposals that invoke increased group size as the selection pressure implicated in the emergence of human musicality are the **coalition signaling** theory (section 2.3), the **credible signaling of parental attention** theory (section 2.4), and certain **therapeutic function** theories, such as the hypothesis that rhythmic synchronization may have emerged as a **coping co-stereotypy** (section 2.5). Mechanistic theories that relate the emergence of musical rhythm to bipedalism, such as the **incidental sounds of locomotion** theory (section 3.1) and the **coping co-stereotypy** hypothesis (section 2.5) are also contradicted by the findings of Schruth et al. (2019), since the emergence of bipedalism is related to terrestrial life (section 3.1). All these theories need, however, not be dismissed given their more specific focus on rhythmic cognition as opposed to musicality in general. In contrast, Schruth et al. (2019) used a composite index of musicality, and thus no separate correlation analyses were conducted between socio-environmental variables and distinct components of musicality. Since this composite index stresses complexity as a fundamental feature of human music, it would also fail to capture several forms of human music in which simplicity predominates. Moreover, it is reasonable to assume that acoustic structural complexity was not one of the foundational features in the emergence of human musicality, but likely it was a culturally evolved one (Nettl, 2000). Thus, theories that run against the findings of Schruth et al. (2019) may still offer good accounts of putative selection pressures involved in the emergence of human rhythmic cognition. Additional research along the lines suggested in this section will be instrumental in clarifying such issues.

The mere presence of certain rhythmic features in the vocalizations (and movements) of other species does not necessarily imply that individuals of these species possess rhythmic abilities that are relevant to the evolution of musicality. In other words, rhythmic structure in animal signals does not entail that conspecifics *perceive* rhythmic structures in these behaviors in the same way as humans can, nor that they will be able to detect rhythmic structure in arbitrary stimuli. Nevertheless, sufficient and systematic data on consistent variations as a function of relevant variables (e.g., contexts, social consequences, sex, potential physiological correlates) can provide valuable insights into the evolution of rhythmic abilities.

A recent playback study with Northern elephant seals (*Mirounga angustirostris*) demonstrates the importance of studying non-melodic calls (Mathevon et al., 2017). Throughout the breeding season, after the male hierarchy is established through fighting, mature males of this species emit threat calls that consist of a rhythmic series of pulses with individual-specific tempo and timbre features. By playing back calls recorded from dominant males, as well as altered versions of these calls, in which tempo, timbre and rhythmic structure were modified, Mathevon et al. (2017) showed that Northern elephant seals can discriminate a variety of tempi and rhythmic patterning. Indeed, the navigational decisions of *beta* males were influenced by stimulus manipulations, in that they would inhibit approach when call features were within the range of individuals higher up in the hierarchy. This finding indicates that Northern elephant seal males memorize and recognize the individual rhythmic signatures of their rivals, and use this information for navigating social life in the colony.

Finally, the usefulness of investigating non-vocal rhythmic behaviors beyond entrainment is illustrated by two recent primate studies: an observational study of cao vit gibbons (*Nomascus nasutus*, Fan et al., 2016) and an experimental study with chimpanzees

(Hattori & Tomonaga, 2019). Female cao vit gibbons were observed making sequential movements of single body parts that have been likened to human so-called ‘robot dance’ and described as rhythmic. Directed towards males, these displays seemingly function as a form of sexual solicitation. The chimpanzee study, on the other hand, brings evidence of rhythmic engagement with an exogenous auditory rhythm, which was documented in both males and females. Both studies provide examples of spontaneous bodily displays with rhythmic structure, as opposed to the trained bodily movements featured by the experimental studies on entrainment reviewed in section 3.1. Non-vocal rhythmic displays have been also documented in palm cockatoos (*Probosciger aterrimus*), where they seem to have a courtship function, in the form of drumming displays that are directed by males towards females (Heinsohn et al., 2017).

Insert Figure 1 about here

In conclusion, after mostly focusing on audiomotor entrainment, comparative research is now beginning to direct attention to other rhythmic abilities. Recent studies have, for example, targeted isochrony detection and beat perception, have attempted to map and compare the rhythmic structure of animal calls to that of musical rhythms, and have begun to examine non-vocal rhythmic behaviors. We argue that this research can further be expanded by systematically targeting rhythmic abilities that, in human ontogeny precede the emergence of audiomotor entrainment (as reviewed above), and structural features of musical rhythm that are cross-culturally universal (see Figure 1 for a synopsis of future research directions outlined in this section). Systematic research targeting such perception and production abilities separately will put us on a better track for determining ancestral traits that are shared with other species, as well as potential innovations that have emerged in the *Homo* lineage.

4.2. Rhythmicity is not modality specific - moving beyond the auditory domain

Rhythmic synchronization is epitomized by the ability to align behaviors with a musical beat. In humans, however, this ability extends beyond the realm of music and dance, and beyond a sensitivity to rhythmic stimuli in the auditory modality. It enables a great variety of activities that rely on precisely timed teamwork through jointly synchronizing to a rhythm, from rowing in a team to military marches. It is highly likely that, throughout human history, rhythmic synchronization has been frequent in coordinating work during routine chores, such as grain pounding, rope pulling, tree chopping, group paddling, etc. (e.g., Dissanayake, 2009a).

Recent experimental evidence suggests that, rather than being modality specific, human rhythmic abilities are susceptible to multi-modal stimulation. Initially, evidence from several studies showed that humans synchronized better with auditory than visual stimuli (for a review see Repp & Su, 2013). These findings appeared to conform to a broader pattern, whereby, unlike other primates, humans were assumed to have superior short-term and recognition memory in the auditory modality compared to other modalities. Inferior memory performance in the auditory modality compared to the visual or tactile modality has been documented in Old World and New World monkeys, as well as one chimpanzee (as reviewed by Bigelow & Poremba, 2014). However, when subjected to experimental testing, the assumption of human superiority in the auditory modality has not been validated. More specifically, human participants, similarly to other primates, show inferior memory for auditory stimuli compared to both visual and tactile stimuli (Bigelow and Poremba 2014).

Recent research also suggests that the auditory advantage, which humans appeared to exhibit in the context of rhythmic synchronization, could be stimulus- and experience-driven (e.g., Gan et al., 2015; Iversen et al., 2015). Indeed, the auditory advantage disappears when humans synchronize to dynamic visual stimuli (e.g., bouncing balls) as opposed to static ones

(e.g., blinking flashes), which were the customary visual stimulus in early synchronization studies (e.g., Gan et al., 2015; Iversen et al., 2015). Moreover, just like audiomotor entrainment (e.g., Madison, 2014), visuomotor entrainment to complex stimuli improves in the presence of metrical subdivisions (Su, 2016), thus extending the evidence on the amodal nature of rhythmic synchronization even to the higher-level of meter processing. Contrary to what might be expected, visual - as opposed to auditory - information has a greater impact on people's ratings of musical performance quality (Griffiths & Reay, 2017; Tsay, 2013). For example, humans are significantly more successful at identifying winners of musical competitions using visual information of their performance only, compared to using auditory information only (Tsay, 2013). Moreover, when presented with manipulated audiovisual recordings of a classical piece in which the audio recording of a professional performer is paired with the visual stream of an amateur performer and vice-versa, both musicians and non-musicians rate the latter (i.e., the visual recording of the professional performer plus audio recording of the amateur performer) as being of higher musical quality (Griffiths & Reay, 2017).

Consistent with these findings, brain imaging evidence shows that auditory, visual and vibrotactile rhythms activate a common neural network for beat detection, which is mostly supramodal, thus transcending sensory modalities (Araneda et al., 2016). In fact, this might extend to all vocal communication (including speech), considering the multisensorial nature of most neocortical areas, including the primary auditory, somatosensory and visual cortices, even though these are, typically, regarded as the foundation of unimodal processing (Ghazanfar, 2009; Ghazanfar et al., 2006). Research shows, for example, that the auditory and visual components of speech are processed simultaneously at the earliest level in the neural circuitry, rather than unimodally and serially (Chandrasekaran et al., 2009; Rosenblum, 2008).

Evidence from early human infancy or prenatal development further supports the

hypothesis that human rhythmic abilities are not modality specific. Research shows that fetuses are able to discriminate rhythmic vibrotactile patterns (e.g., the mother's walking vs. swaying vs. rocking), as evidenced by changes in heart-rate response patterns (Lecanuet & Jacquet, 2002). Such vibrotactile stimulation appears to be crucial for normal fetal development, which is also evidenced by the benefits of rocking in promoting typical development in prematurely born infants (as reviewed by Provasi et al., 2014), as well as by the worldwide use of rocking to calm young infants (Richter & Ostovar, 2016). In both fetuses and young infants, cross-modal (e.g., vibro-acoustic, audio-visual) presentation of a rhythm provides a processing advantage compared to unimodal presentation of an auditory rhythm. As such, fetal responses are higher when fetuses are exposed to vibro-acoustic stimuli compared to unimodal stimuli (Kisilevsky & Muir, 1991; Kisilevsky et al., 1992). In addition, three- and five-month old infants can learn to discriminate a rhythmic pattern when the pattern is presented synchronously in the visual and auditory modalities, but fail to learn the rhythm when it is presented unimodally, in either the visual or auditory modality (Bahrick & Lickliter, 2000; Bahrick et al., 2002).

Taken together, the findings presented in this section suggest that unimodal theories likely provide an incomplete account of musicality in general, and of rhythmic abilities in particular. Examining the rhythmic behaviors of other species beyond vocal production and auditory reception will therefore likely provide further insights into the evolution of human musicality. It is worth noting that a similar need exists in human developmental research, where the overwhelming majority of studies focus on rhythms in vocal communication and auditory perception, while the perception and production of *bodily* rhythms remain relatively underexplored (Provasi et al., 2014). For the purposes of comparative research, relevant (non-vocal or cross-modal) behaviors are those that fulfil a communicative function in other species, thus excluding biological rhythms such as breathing or locomotion. Observational

and playback experiments similar to those mentioned in section 4.1. would be instrumental in this respect, but with a focus informed by research questions pertaining to modality. Investigating rhythmic communicative behaviors that take the form of bodily displays may reveal behavioral patterns and rhythmic features that are comparable to e.g., human dance and musical rhythm. For example, the detailed analysis of rhythmic co-locomotion in chimpanzees (i.e., the so-called ‘Conga line’) has recently led to the proposal that human proto-dance may have been rooted in stress-relieving behaviors that initially did not involve any sound (Lameira et al., 2019). Our own observations show that, generally, chimpanzees exhibit several rhythmic behaviors that seem to accomplish a communicative function, as these behaviors are primarily exhibited in a social context (Sauciuc et al., 2019a), and exert a social influence on conspecifics (Sauciuc et al., 2019b). Many parrot species also show spontaneous bodily movements (e.g., head bobbing) that seem to have a communicative function.

Interestingly, and as noted above, among avian vocal learners, evidence of audiomotor entrainment has only been found in parrots. In parrot species, this ability might be expressed spontaneously and is potentially linked to patterns of brain organization that, among avian species, are specific to psittaciformes. In particular, while vocal motor pathways are anatomically adjacent to the auditory pathways in songbirds, they are more distant in parrots (Jarvis & Mello, 2000). Moreover, psittaciforme birds possess an additional vocal system that also includes non-vocal motor pathways (Chakraborty et al., 2015), which are active during e.g., hopping and head bobbing movements (Feeders et al., 2008). Based on these findings, from both chimpanzees and parrots, and also considering that, currently, spontaneous entrainment has only been reported for chimpanzees and larger parrots, it is possible that rhythmic gestural/bodily communication (especially if exhibiting a certain level of flexibility and voluntary control) represents a prerequisite trait for spontaneous rhythmic

1595 synchronization to emerge.

1596 Non-communicative rhythmic behaviors (e.g., respiration, locomotion) are not
 1597 completely without relevance for the evolutionary study of musical rhythm. Comparing e.g.,
 1598 respiratory and/or locomotor rates with the output rates of communicative rhythmic behaviors
 1599 in various modalities will provide valuable data to determine if there are biological couplings
 1600 across modalities and behaviors, which might point to underlying mechanisms in the form of,
 1601 e.g., species-specific internal-clock like mechanisms. This could also be informative for
 1602 understanding how and why optimal rhythmic rates emerge, and for explaining variations in
 1603 such rates. For example, the optimal finger tapping rate in human adults and common
 1604 chimpanzees appears to be similar (around 500-600 ms inter-onset-interval), while the rate of
 1605 drum striking in bonobos and human children are typically much lower (around 300-450 ms
 1606 interval). Such data will also be helpful for assessing the predictions of the **acoustic**
 1607 **advantages** hypothesis (Larsson & Abbott, 2018), which links the emergence of rhythmic
 1608 synchronization to locomotor-respiratory couplings present in fish, and therefrom to
 1609 audiomotor couplings between forepaw processing and vocalizing in tetrapods. Biological
 1610 constraints grounded on such couplings appear to drive the rhythm of isochronous
 1611 echolocation calls in several bat species, where such calls are coupled with wingbeat and
 1612 breathing cycles (for a review see Burchardt et al., 2019). Interestingly, however, when
 1613 sensory needs are higher - such as during prey approach or when navigating complex
 1614 environments - the rates of echolocation calls are dissociated from wingbeat and respiratory
 1615 rates. Recent evidence further shows that in the Greater sac-winged bat (*Saccopteryx*
 1616 *bilineata*), a single frequency dominates the rhythm of both (non-vocal) echolocation calls
 1617 and (vocal) social calls, such as pup isolation calls and male territorial songs (Burchardt et al.,
 1618 2019). It is intriguing that these types of calls share a similar frequency (which is also shared
 1619 with non-communicative rhythms), considering that isolation calls are innate (and later in life

repurposed as appeasement calls in males) while territorial songs, in contrast, are learned (Eckenweber & Knörnschild, 2013).

Systematically mapping the bio-physical properties of animal gaits and the rhythmic patterns they produce will allow testing hypotheses that emphasize **bipedalism** as a crucial trait for the emergence of human rhythmic abilities. As reviewed in section 3.1, it has been proposed that bipedal walking more readily affords rhythmic synchronization (compared with quadrupedal locomotion), by virtue of entailing biphasic isochronous movement (see Larsson, 2013; Larsson et al., 2019). Testing this hypothesis requires the study of locomotion rates and perceptual rhythmic abilities in species in which locomotion is accomplished through biphasic periodic movements, such as flight and walking in birds, or stotting (i.e., springing with all four legs in the air) in mammalian species. Also, it remains to be investigated if other gait patterns have an influence on rhythm *perception*, for example, if quadrupeds are primed by quadrupedal gait patterns, thus showing a sensitivity or preference for non-binary rhythms.

Finally, experimental studies with relevant nonhuman species should also compare rhythmic production and perception across modalities, as well as compare rhythmic vs. non-rhythmic perception in different modalities. Such experimental data will be helpful in assessing hypotheses that rhythmic signaling entails transmission advantages and patterning flexibility to facilitate the recognition of species-specific calls (as reviewed in section 2.1). Recognition and mnemonic advantages are also predicted by **the dynamic attending theory** (as reviewed in 3.1). While memory for auditory stimuli poses a challenge for many species, including humans (Bigelow & Poremba, 2014 and references therein), this challenge may be lessened by rhythmic patterning, as e.g., suggested by several of the theoretical proposals reviewed in sections 2 and 3. Initial studies have led to the assumption that humans entrain better to auditory compared to visual stimuli, while other primates show the opposite pattern. Recent studies, however, have begun to challenge these assumptions. For example, there is

evidence that humans entrain equally well to visual and auditory stimuli when presented with appropriate stimuli (Gan et al., 2015; Iversen et al., 2015). In addition, flexible entrainment to auditory stimuli can be trained in macaques to levels that match those for entrainment to visual stimuli (Gámez et al., 2018; Takeya et al., 2017). Moreover, preliminary data indicate that chimpanzees who have been trained to tap rhythmically display similar levels of rhythmic coordination with a conspecific, regardless of whether feedback is given in the auditory channel only or in both the auditory and visual channels (Yu & Tomonaga, 2018).

Insert Figure 2 about here

Summing up, our call to expand the focus of research on the rhythmic abilities of non-human species beyond auditory perception and vocal production encompasses behaviors with - as well as without - a *communicative* function (as also shown in Figure 2). Such research should examine the presence, range and function of rhythmic behaviors across modalities of perception and production, as well as the flexibility of such behaviors and the degree of intentional control over them. Within species comparisons of rhythmic abilities exhibited across modalities and cross-species comparisons focusing on single modalities will allow us to test long-held assumptions (e.g., human auditory superiority) and more recent hypotheses (e.g., the beacon effect hypothesis, the signal conservation hypothesis, the incidental sounds of locomotion hypothesis, the acoustic advantage hypothesis) on the nature of musical rhythms and its underlying mechanisms.

4.3. A diversity of functional contexts - moving beyond courtship and territorial defense

As reviewed in section 2, rhythmic behavior has been associated with both individual and social benefits in humans, yet it remains unclear which of these have an ancestral

relevance for the evolution of musicality. Interestingly, human music appears to exhibit intriguing form-function regularities, as revealed by a recent cross-cultural study. Independent of their cultural background, people reliably match lullabies, dance, love, and healing songs from across the world with their functional context (Mehr et al., 2018). Such functional categorization is partly driven by rhythmic features related to emotional valence, or levels of arousal conveyed by a given musical piece. For example, dance songs are characterized by higher rhythmic complexity, a steadier beat and a higher level of arousal and pleasantness compared to lullabies. This suggests the presence of a systematic relationship between musical form and its affective effects and functional uses, with potentially ancestral origins that are rooted in behavioral (vocal, gestural) expressions of emotion.

The comparative literature reviewed above suggests that rhythmic structure may be common in animal signals. Preliminary evidence further shows that the mammalian brain automatically attunes its oscillations to exogenous rhythms (e.g., Noda et al., 2017) and detects on-beat emphasis (Rajendran et al., 2020), which suggests a preparedness for rhythm detection, whose functional relevance, however, remains elusive. A systematic study of the rhythmic behaviors exhibited by a diversity of species, and the functional contexts in which they occur (as summarized in Figure 3), is thus likely to provide valuable data both for understanding the evolution of human musicality, as well as the species-specific functions of various forms of rhythmic behaviors, and whether functional trends can be delineated across species. This research should also consider modalities of perception (e.g., auditory, visual, as outlined in 4.2.) and the full spectrum of rhythmic abilities (as outlined in 4.1), and provide detailed descriptions of the structure of rhythmic behavior, as a crucial first step in understanding their ecology and evolution (for a similar point see Caselli et al., 2014).

Generally, not only is there a lack of studies focused on the potential significance of rhythmic signals beyond the contexts of mating and territorial defense (as reviewed in 2.1),

but there is also a bias towards studying rhythmically structured behavior in the vocal channel only (as reviewed in 4.2). A handful of studies suggests, however, that rhythmic signals may occur in other contexts as well, although these focus on vocal behaviors. For example, rhythmic vocalizations with a socialization function have been described in gelada monkeys (*Theropithecus gelada*: Richman, 1978) and in long-finned pilot whales (*Globicephala melas*: Zwamborn & Whitehead, 2016). The geladas appear to be able to synchronize the tempo and rhythm of their own vocalizations with the vocalizations emitted by a conspecific, while the pilot whales emit their rhythmic vocalizations primarily when gathered in larger groups at sea surface.

Insert Figure 3 about here

Rhythmic vocalizations in aggressive contexts have been described, for example, in bats and dolphins. In the greater sac-winged bat, these take the form of male-emitted appeasement signals (Fernandez & Knörnschild, 2017). In aggressive contexts, Atlantic spotted dolphins (*Stenella frontalis*) and bottlenose dolphins (*Tursiops truncatus*) exhibit two, respectively three, distinct rhythmic vocalizations. Dolphins emit these calls communally, during both intra- and interspecific aggressive encounters, where they also perform bodily displays, although the temporal structure of these displays remains to be studied in detail (Herzing & Dolphin, 2015).

Recently, our own research group set out to examine the presence, range, form and functional significance of rhythmic behaviors in chimpanzees. Our preliminary reports suggest that rhythmic behaviors are frequent in chimpanzees, and may occur in a variety of contexts, including initiation of social interaction, social play, solitary play, dominance display, territorial group display, affiliative interaction, and courtship (Sauciuc et al., 2019a).

1719 We documented considerable flexibility with respect to form and function, as these behaviors
 1720 can be vocal or gestural, can take on a diversity of forms, involving various effectors (whole
 1721 body, head, hand, arm, foot, or combinations of these), and, generally, are not rigidly bound
 1722 to a specific context (Sauciuc et al., 2019b). Rhythmic bouncing, for example, can occur as a
 1723 signal for initiation of social interaction, as well as during social play, and during courtship.
 1724 To date, such diversity and functional flexibility of rhythmic behaviors has only been
 1725 documented in one other species: humans. It is not clear, however, if chimpanzees are indeed
 1726 exceptional in this respect, or if this is the result of a comprehensive lack of data from other
 1727 species. This further highlights the significance of the point highlighted in this section, i.e.,
 1728 the need for the systematic study of the rhythmic behaviors exhibited by a diversity of
 1729 species, and the functional contexts in which they occur.

1730 In chimpanzees, rhythmic behaviors have previously been associated with stress-
 1731 induced stereotypical movement, and, thus, with high levels of arousal (Chamove, 1989;
 1732 Lopresti-Goodman et al., 2013). In the majority of the cases that we documented, piloerection
 1733 was absent, thus suggesting that these rhythmic behaviors were not a mere expression of
 1734 arousal. In addition, rhythmic behaviors co-occurred frequently with the relaxed open-mouth
 1735 face, which is a correlate of states of relaxation and positive emotion (Parr & Waller, 2006).
 1736 In chimpanzees, both males and females engaged in such rhythmic behaviors, while for the
 1737 great majority of other species where rhythmic output has been described - almost exclusively
 1738 in the form of vocalizations - this phenomenon is restricted to males (as reviewed in sections
 1739 **2.1.** and **4.1**). Based on extant data, the vocalizations of some primate species - such as titi
 1740 monkeys and tarsiers - could be an exception to this male bias. While we found that male
 1741 chimpanzees performed rhythmic behaviors more often than female chimpanzees, the social
 1742 efficacy of these behaviors did not seem to differ between sexes, displaying similar levels of
 1743 social influence on conspecifics (Sauciuc et al., 2019b).

1744

1745 **5. Conclusions: The fallacy of “one and only-theories”**

1746 As reviewed in sections 2 and 3, a comprehensive bias towards theoretical work - and
 1747 a paucity of empirical work - dominates the evolutionary study of musicality. During the last
 1748 two decades, there has been an explosion of theoretical proposals aimed at explaining why
 1749 and how humans have evolved into musical beings. Most often, these proposals on putative
 1750 evolutionary scenarios focus on pinpointing *the one* function or *the one* functional context
 1751 that had precedence in ancestral times, thus being the primary driving force in the evolution of
 1752 musicality. Such endeavors, however, are invariably predestined to result in a chicken-or-the-
 1753 egg debate, as complex traits, such as musicality, are inherently compositional, and thus, their
 1754 evolution is irreducible to the evolution of a single component with a putatively single
 1755 adaptive function. Even when looking at single components, however, multiple factors
 1756 combine into making something adaptive, in the sense that traits and behaviors emerge
 1757 ontogenetically from gene-brain-body-environment feedback loops, and it is therefore
 1758 impossible to propose a ‘one-and-only’ origin story behind any evolved thing. This is perhaps
 1759 best resolved by not asking “what has musicality evolved for”, which easily leads to the idea
 1760 of a singular reason, but “what has made musicality adaptive”, which can be many things.

1761 The review of proximate mechanisms hypothesized to be involved in the perception
 1762 and production of musical rhythm (provided in 3.2) demonstrates this point, suggesting that
 1763 musicality has emerged in concert with, and not independently of, other communication
 1764 systems as well as a plethora of socio-cognitive abilities. Some of these appear to be
 1765 widespread in the animal kingdom (e.g., predictive coding, neural resonance, emotional
 1766 contagion), others appear to be human- (or Hominid-) specific (e.g., shared intentionality) or
 1767 present in a few highly intelligent species (e.g., episodic foresight). How these prerequisites
 1768 for evolving music-like rhythmic behaviors relate to one another to contribute to the

1769 emergence of musical rhythm is still to be investigated.

1770 With the recent advent of compositional approaches to musicality, the empirical
 1771 research targeting hypotheses regarding the evolution of musicality has gained some
 1772 momentum. However, compared to the number of theoretical proposals published in the last
 1773 two decades, the number of empirical studies continues to remain disproportionately low. In
 1774 this paper, we have specifically focused on theories related to the rhythmic component of
 1775 musicality, and on abilities regarded as prerequisites for perceiving and producing rhythmic
 1776 structures similar to those encountered in music. Currently, the findings reported in these
 1777 studies are far from providing a coherent picture of the evolution of musicality and, in
 1778 particular, the rhythmical abilities that make it possible. They succeed, nevertheless, at
 1779 revealing that even the most focused and seemingly straightforward questions beg highly
 1780 complex answers. Since the evolutionary study of musicality is inherently multidisciplinary,
 1781 seeking answers to these questions will require interdisciplinary collaboration, and research
 1782 that will both *zoom in* on specific mechanisms, contexts and species, and *zoom out* from
 1783 detailed analyses to broader views.

1784 The review in sections 2 and 3 is intended to provide a background for a research
 1785 agenda centered on topics that have been so far neglected in the comparative study of
 1786 rhythmic abilities, but that are highly relevant for understanding where musicality comes
 1787 from. In particular, we call for a widened focus of research on rhythmic cognition in non-
 1788 human species, so as to include additional rhythmic abilities besides entrainment (4.1), other
 1789 modalities of rhythmic expression besides the vocal modality (4.2), and a systematic focus on
 1790 the functional contexts in which rhythmic abilities are spontaneously employed (4.3).
 1791 Broadening the scope of comparative research to include rhythmic abilities (e.g., rhythm
 1792 engagement, isochrony detection, beat perception) that are considered precursors to accurate
 1793 entrainment (at least in human ontogeny) is necessary if the purpose is to track the

1794 evolutionary roots of advanced rhythmic abilities. Expanding research to other modalities
 1795 than auditory perception and vocal communication is also important. In particular, there is a
 1796 need for systematic research on the perception and production of bodily rhythms, since
 1797 rhythmic movement is an essential feature of musicality, and given suggestions that
 1798 musicality may have originated as bodily-coupled entrainment in the absence of an auditory
 1799 stimulus. Finally, systematically mapping the contexts in which animal rhythmic signals
 1800 occur spontaneously is crucial to understanding the functional significance of rhythmic
 1801 communication, its efficacy and the factors that underlie and influence it. The insights gained
 1802 from such research will, in turn, inform theories about the adaptive significance of musical
 1803 rhythm. In this context, special attention should be given to rhythmic communication through
 1804 bodily movement, as this is conducive to rhythmic engagement, which likely promotes and
 1805 reaffirms social affiliation.

1806 Extant comparative data suggests that most of the mechanisms that enable musical
 1807 rhythm may not be musicality specific. Some of these mechanisms appear to be widespread
 1808 across species, and, thus, phylogenetically old. Yet, detailed analyses and cross-species
 1809 comparisons are needed to determine the extent to which the neural networks and
 1810 psychobiological mechanisms that mediate the perception, production and appreciation of
 1811 musical rhythm in humans have identical counterparts in other species. A few high-level
 1812 cognitive and socio-cognitive abilities that are relevant to the evolution of musical rhythm
 1813 (e.g., relational processing, causal reasoning, enhanced memory and foresight, shared
 1814 intentionality, imitation and pantomime) appear to be rare in the animal kingdom, being
 1815 exclusive to highly intelligent species, such as corvids and hominids, or restricted to hominids
 1816 only. The emergence of these high-level abilities is generally related to adaptive pressures
 1817 posed by complex social organization and complex foraging strategies. Seemingly, however,
 1818 at some point in the evolution of our species, these abilities have been co-opted into

1819 rhythmical, communal displays. Additional traits that, so far, do not seem to have a
 1820 counterpart in the animal kingdom, and which are related to human creativity, narrativity,
 1821 aesthetics and linguistic communication, have further shaped such behaviors into what we
 1822 today call music and dance.

1823

1824 **Author contributions**

1825 Gabriela-Alina Sauciuc has conceived and drafted the paper. Tomas Persson and Elainie
 1826 Madsen have critically reviewed it.

1827

1828 **Funding**

1829 This work has been supported by grant 2017:0942:1 from Riksbankens Jubileumsfond (The
 1830 Bank of Sweden Tercentenary Foundation).

1831 **Conflicts of interests**

1832 The authors have no conflicts of interest to declare that are relevant to the content of this
 1833 article.

1834 **Ethical statement.** Not applicable.

1835 **Data availability statement.** Not applicable.

1836

1837 **References**

- 1838 Aalbers, S., Fusar-Poli, L., Freeman, R., Spreen, M., Ket, J.C., Vink, A., Maratos, A.,
 1839 Crawford, M., Chen, X., & Gold, C. (2017). Music therapy for depression. *The Cochrane*
 1840 *database of systematic reviews*, 11, CD004517.
- 1841 Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*,
 1842 162, 46-59.
- 1843 Alcaro, A., Huber, R., & Panksepp, J. (2007). Behavioral functions of the mesolimbic

- 1844 dopaminergic system: An affective neuroethological perspective. *Brain Research Reviews*,
 1845 56, 283-321.
- 1846 Alward, B.A., Rouse, M.L., Balthazart, J., & Ball, G. (2017). Testosterone regulates birdsong
 1847 in an anatomically specific manner. *Animal Behaviour*, 124, 291-298.
- 1848 Alworth, L., & Buerkle, S.C. (2013). The effects of music on animal physiology, behavior
 1849 and welfare. *Lab Animal*, 42, 54-61.
- 1850 Andrews, P.W., Gangestad, S., & Matthews, D. (2002). Adaptationism – how to carry out an
 1851 exaptationist program. *Behavioral and Brain Sciences*, 25, 489 - 504.
- 1852 Anshel, A., & Kipper, D. A. (1988). The influence of group singing on trust and cooperation.
 1853 *Journal of Music Therapy*, 25(3), 145–155. <https://doi.org/10.1093/jmt/25.3.145>.
- 1854 Arom, S. (2000). Prolegomena to a biomusicology. In N. L. Wallin, B. Merker, & S. Brown
 1855 (Eds.), *The Origins of music* (pp. 27–29). Cambridge, Mass: The MIT Press.
- 1856 Bachner-Melman, R., Dina, C., Zohar, A., Constantini, N., Lerer, E., Hoch, S., Sella, S.,
 1857 Nemanov, L., Gritsenko, I., Lichtenberg, P., Granot, R., & Ebstein, R. (2005). AVPR1a and
 1858 SLC6A4 Gene Polymorphisms Are Associated with Creative Dance Performance. *PLoS*
 1859 *Genetics*, 1.
- 1860 Bachorowski, J.-A., & Owren, M. J. (2003). *Sounds of Emotion: Production and Perception*
 1861 *of Affect-Related Vocal Acoustics*. In P. Ekman, J. J. Campos, R. J. Davidson, & F. B. M. de
 1862 Waal (Eds.), *Annals of the New York Academy of Sciences: Vol. 1000. Emotions inside out:*
 1863 *130 years after Darwin's: The expression of the emotions in man and animals* (p. 244–265).
 1864 New York Academy of Sciences.
- 1865 Bahrick, L., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates
 1866 discrimination of tempo in 3-month-old infants. *Developmental psychobiology*, 41 4, 352-63.
- 1867 Bahrick, L., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and
 1868 perceptual learning in infancy. *Developmental psychology*, 36 2, 190-201.

- 1869 Baskerville, T. A., & Douglas, A. J. (2010). Dopamine and oxytocin interactions underlying
1870 behaviors: Potential contributions to behavioral disorders. *CNS Neuroscience & Therapeutics*,
1871 *16*(3), e92–e123.
- 1872 Bigelow, J., & Poremba, A. (2014). Achilles' Ear? Inferior Human Short-Term and
1873 Recognition Memory in the Auditory Modality. *PLoS ONE*, *9*, e89914.
- 1874 Blood, A., & Zatorre, R. (2001). Intensely pleasurable responses to music correlate with
1875 activity in brain regions implicated in reward and emotion. *Proceedings of the National*
1876 *Academy of Sciences of the United States of America*, *98*, 11818 - 11823.
- 1877 Bouwer, F.L., Zuijlen, T.V., & Honing, H. (2014). Beat Processing Is Pre-Attentive for
1878 Metrically Simple Rhythms with Clear Accents: An ERP Study. *PLoS ONE*, *9*, e97467.
- 1879 Brandt, P.A. (2009). Music and how we became human: a view from cognitive semiotics -
1880 exploring imaginative hypotheses. In S. Malloch & C. Trevarthen (Eds.), *Communicative*
1881 *musicality: Exploring the basis of human companionship* (pp. 31–44). Oxford University
1882 Press.
- 1883 Brown, S. (2000a). Evolutionary models of music: From sexual selection to group selection.
1884 In Tonneau, F., Thompson, N. S. (eds.), *Perspectives in ethology*. 13: Behavior, evolution and
1885 culture (pp. 231–81). New York: Plenum Publishers.
- 1886 Brown, S. (2000b). "The 'musilanguage' model of music evolution," in *The Origins of Music*,
1887 eds N. L. Wallin, B. Merker and S. Brown (Cambridge, MA: MIT Press), 271–300.
- 1888 Brown, S. (2007). Contagious heterophony: A new theory about the origins of music.
1889 *Musicae Scientiae*, *11*, 26 - 3.
- 1890 Brown S. (2017). A Joint Prosodic Origin of Language and Music. *Frontiers in psychology*, *8*,
1891 1894.
- 1892 Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, *41*,
1893 229 - 248.

- 1894 Brumm, H., & Slater, P. (2006). Ambient noise, motor fatigue, and serial redundancy in
1895 chaffinch song. *Behavioral Ecology and Sociobiology*, 60, 475-481.
- 1896 Bryant, G. (2013). Animal signals and emotion in music: coordinating affect across groups.
1897 *Frontiers in Psychology*, 4.
- 1898 Burchardt LS, Norton P, Behr O, Scharff C, Knörnschild M. 2019 General isochronous
1899 rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *R. Soc.*
1900 *open sci.* 6: 181076.
- 1901 Carson, D., Guastella, A., Taylor, E., & McGregor, I. (2013). A brief history of oxytocin and
1902 its role in modulating psychostimulant effects. *Journal of Psychopharmacology*, 27, 231 -
1903 247.
- 1904 Caselli, C.B., Mennill, D., Bicca-Marques, J., & Setz, E. (2014). Vocal behavior of black-
1905 fronted titi monkeys (*Callicebus nigrifrons*): Acoustic properties and behavioral contexts of
1906 loud calls. *American Journal of Primatology*, 76.
- 1907 Cate, C.T., & Spierings, M.J. (2019). Rules, rhythm and grouping: auditory pattern perception
1908 by birds. *Animal Behaviour*, 151, 249-257.
- 1909 Cate, C.T., Spierings, M., Hubert, J., & Honing, H. (2016). Can Birds Perceive Rhythmic
1910 Patterns? A Review and Experiments on a Songbird and a Parrot Species. *Frontiers in*
1911 *Psychology*, 7.
- 1912 Celma-Miralles, A., & Toro, J. M. (2020a). Discrimination of temporal regularity in rats
1913 (*Rattus norvegicus*) and humans (*Homo sapiens*). *Journal of Comparative Psychology*,
1914 134(1), 3–10.
- 1915 Celma-Miralles, A., & Toro, J.M. (2020b). Non-human animals detect the rhythmic structure
1916 of a familiar tune. *Psychonomic bulletin & review*. 27, 694–699.
- 1917 Chakraborty, M., Walløe, S., Nedergaard, S., Fridel, E.E., Dabelsteen, T., Pakkenberg, B.,
1918 Bertelsen, M., Dorrestein, G., Brauth, S., Durand, S.E., & Jarvis, E. (2015). Core and Shell

- 1919 Song Systems Unique to the Parrot Brain. *PLoS ONE*, 10, e0118496.
- 1920 Chamove, A. (1989). Environmental Enrichment: A Review. *Animal Technology*, 40, 3.
- 1921 Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A.A. (2009).
- 1922 The Natural Statistics of Audiovisual Speech. *PLoS Computational Biology*, 5, e1000436.
- 1923 Chang, S., Brent, L., Adams, G.K., Klein, J., Pearson, J., Watson, K.K., & Platt, M. (2013).
- 1924 Neuroethology of primate social behavior. *Proceedings of the National Academy of Sciences*,
- 1925 110, 10387 - 10394.
- 1926 Chaudhry S.R., Bhimji S.S. (2018). Biochemistry, Endorphin. StatPearls. StatPearls
- 1927 Publishing. PMID 29262177. Retrieved 20 February 2019.
- 1928 Chu, K., & Harcourt, P. (2004). Behavioral correlations with aberrant patterns in humpback
- 1929 whale songs. *Behavioral Ecology and Sociobiology*, 19, 309-312.
- 1930 Cirelli, L.K., Einarson, K.M., & Trainor, L. (2014). Interpersonal synchrony increases
- 1931 prosocial behavior in infants. *Developmental science*, 17 6, 1003-11.
- 1932 Cirelli, L.K., Wan, S.J., Spinelli, C., & Trainor, L. (2017). Effects of Interpersonal Movement
- 1933 Synchrony on Infant Helping Behaviors: Is Music Necessary? *Music Perception: An*
- 1934 *Interdisciplinary Journal*, 34, 319-326.
- 1935 Clink, D., Tasirin, J.S., & Klinck, H. (2020). Vocal individuality and rhythm in male and
- 1936 female duet contributions of a nonhuman primate. *Current Zoology*, 66, 173 - 186.
- 1937 Cohen, E., Ejsmond-Frey, R., Knight, N., & Dunbar, R. (2009). Rowers' high: behavioural
- 1938 synchrony is correlated with elevated pain thresholds. *Biology Letters*, 6, 106 - 108.
- 1939 Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (*Zalophus*
- 1940 *californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non
- 1941 vocal mimic. *Journal of comparative psychology*, 127 4, 412-27.
- 1942 Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild Chimpanzees Produce
- 1943 Group-Specific Calls: a Case for Vocal Learning? *Ethology*, 110, 221-243.

- 1944 Crockford, C., Wittig, R. M., Langergraber, K., Ziegler, T. E., Zuberbühler, K., Deschner, T.
 1945 Urinary oxytocin and social bonding in related and unrelated wild chimpanzees. *Proceedings*
 1946 *of the Royal Society B: Biological Sciences*, 2013; 280 (1755): 20122765.
- 1947 Crockford, C., Wittig, R., & Zuberbühler, K. (2014). An intentional vocalization draws
 1948 others' attention: A playback experiment with wild chimpanzees. *Animal Cognition*, 18, 581-
 1949 591.
- 1950 Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of*
 1951 *Sciences*, 930, 28-42.
- 1952 Cross, I. (2003). Music, cognition, culture and evolution. In I. Peretz & R. J. Zatorre (Eds.),
 1953 *The cognitive neuroscience of music* (pp. 42–56). Oxford: Oxford University Press.
- 1954 Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: J. Murray.
- 1955 Davila-Ross, M., Allcock, B., Thomas, C., & Bard, K. (2011). Apeing expressions?
 1956 Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion*, 11
 1957 5, 1013-20.
- 1958 Dissanayake, E. (2000). *Antecedents of the temporal arts in early mother–infant interaction*.
 1959 In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (p. 389–410). The MIT
 1960 Press.
- 1961 Dissanayake, E. (2009a). *Root, leaf, blossom, or bole: Concerning the origin and adaptive*
 1962 *function of music*. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality:*
 1963 *Exploring the basis of human companionship* (p. 17–30). Oxford University Press.
- 1964 Dissanayake, E. (2009b). *Bodies swayed to music: The temporal arts as integral to*
 1965 *ceremonial ritual*. In S. Malloch & C. Trevarthen (Eds.), *Communicative musicality:*
 1966 *Exploring the basis of human companionship* (p. 533–544). Oxford University Press.
- 1967 Donnet, S., Bartolo, R., Fernandes, J., Cunha, J., Prado, L., & Merchant, H. (2014). Monkeys
 1968 time their pauses of movement and not their movement-kinematics during a synchronization-

- 1969 continuation rhythmic task. *Journal of neurophysiology*, 111 10, 2138-49.
- 1970 Dunbar, R. I. M. (2004). Language, music, and laughter in evolutionary perspective (D. K.
- 1971 Oller & U. Griebel, Eds.; pp. 257–274). MIT Press.
- 1972 Dunbar, R. I. M. (2012). “On the evolutionary function of song and dance,” in *Music,*
- 1973 *Language and Human Evolution*, ed. N. Bannan (Oxford: Oxford University Press), 201–214.
- 1974 doi: 10.1093/acprof:osobl/9780199227341.003.0008.
- 1975 Dunbar, R., Kaskatis, K., Macdonald, I., & Barra, V. (2012). Performance of Music Elevates
- 1976 Pain Threshold and Positive Affect: Implications for the Evolutionary Function of Music.
- 1977 *Evolutionary Psychology*, 10.
- 1978 Eckenweber, M., & Knörnschild, M. (2013). Social influences on territorial signaling in male
- 1979 greater sac-winged bats. *Behavioral Ecology and Sociobiology*, 67, 639-648.
- 1980 Edelman, L. L., & Harring, K. E. (2015). Music and social bonding: The role of non-diegetic
- 1981 music and synchrony on perceptions of videotaped walkers. *Current Psychology: A Journal*
- 1982 *for Diverse Perspectives on Diverse Psychological Issues*, 34(4), 613–620.
- 1983 <https://doi.org/10.1007/s12144-014-9273-y>.
- 1984 Eriksen, N., Miller, L., Tougaard, J., & Helweg, D. (2005). Cultural change in the songs of
- 1985 humpback whales (*Megaptera novaeangliae*) from Tonga. *Behaviour*, 142, 305-328.
- 1986 Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? *The*
- 1987 *Behavioral and brain sciences*, 27 4, 491-503; discussion 503-83.
- 1988 Fan, P., Ma, C., Garber, P., Zhang, W., Fei, H., & Xiao, W. (2016). Rhythmic displays of
- 1989 female gibbons offer insight into the origin of dance. *Scientific Reports*, 6, 34606.
- 1990 Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013). Pant hoot chorusing
- 1991 and social bonds in male chimpanzees. *Animal Behaviour*, 86(1), 189–196.
- 1992 Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., Wada, K.,
- 1993 Mouritsen, H., & Jarvis, E. (2008). Molecular Mapping of Movement-Associated Areas in the

- 1994 Avian Brain: A Motor Theory for Vocal Learning Origin. *PLoS ONE*, 3, e1768.
- 1995 Feldman, R. (2007). Parent–Infant Synchrony. *Current Directions in Psychological Science*,
- 1996 16, 340 - 345.
- 1997 Ferreri, L., Mas-Herrero, E., Zatorre, R., Ripollés, P., Gómez-Andrés, A., Alicart, H., Olivé,
- 1998 G., Marco-Pallarés, J., Antonijoan, R., Valle, M., Riba, J., & Rodríguez-Fornells, A. (2019).
- 1999 Dopamine modulates the reward experiences elicited by music. *Proceedings of the National*
- 2000 *Academy of Sciences*, 116, 3793 - 3798.
- 2001 Fessler, D. M. T., & Holbrook, C. (2016). Synchronized behavior increases assessments of
- 2002 the formidability and cohesion of coalitions. *Evolution and Human Behavior*, 37(6), 502–509.
- 2003 Fitch W.T. 2006. The biology and evolution of music: A comparative perspective. *Cognition*
- 2004 100, 173–215.
- 2005 Fitch, T. 2011. The biology and evolution of rhythm: unravelling a paradox. In, P. Rebuschat,
- 2006 M. Rohmeier, J. A. Hawkins, & I. Cross (Eds.), *Language and Music as Cognitive Systems* (p.
- 2007 73-95). Oxford University Press.
- 2008 Fitch, W. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse
- 2009 perception. *Frontiers in Systems Neuroscience*, 7.
- 2010 Fitch, W. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal*
- 2011 *Society B: Biological Sciences*, 370.
- 2012 Fitch, W. (2016). Dance, Music, Meter and Groove: A Forgotten Partnership. *Frontiers in*
- 2013 *Human Neuroscience*, 10.
- 2014 Frazer, L.N., & Mercado, E. (2000). A sonar model for humpback whale song. *IEEE Journal*
- 2015 *of Oceanic Engineering*, 25, 160-182.
- 2016 Freeman, W. (2000). *A neurobiological role of music in social bonding*. In N. L. Wallin, B.
- 2017 Merker, & S. Brown (Eds.), *The origins of music* (p. 411–424). The MIT Press.
- 2018 Fujii, S., Watanabe, H., Oohashi, H., Hirashima, M., Nozaki, D., & Taga, G. (2014).

- 2019 Precursors of Dancing and Singing to Music in Three- to Four-Months-Old Infants. *PLoS*
 2020 *ONE*, 9, e103192.
- 2021 Gamba, M., Valente, D., Gregorio, C.D., Zanolli, A., Randrianarison, R., Torti, V., &
 2022 Giacomini, C. (2019). Song structure and sex specific features in the indris. *Journal of the*
 2023 *Acoustical Society of America*, 145, 1919-1919.
- 2024 Gámez, J., Yc, K., Ayala, Y.A., Dotov, D., Prado, L., & Merchant, H. (2018). Predictive
 2025 rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate.
 2026 *Annals of the New York Academy of Sciences*, 1423.
- 2027 Gan, L., Huang, Y., Zhou, L., Qian, C., & Wu, X. (2015). Synchronization to a bouncing ball
 2028 with a realistic motion trajectory. *Scientific Reports*, 5, 11974.
- 2029 Gebauer, L., Witek, M., Hansen, N.C., Thomas, J., Konvalinka, I., & Vuust, P. (2016).
 2030 Oxytocin improves synchronisation in leader-follower interaction. *Scientific Reports*, 6,
 2031 38416.
- 2032 Geissmann, T. (2000). *Gibbon songs and human music from an evolutionary perspective*. In
 2033 N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (p. 103–123). The MIT
 2034 Press.
- 2035 Ghazanfar, A.A., & Schroeder, C. (2006). Is neocortex essentially multisensory? *Trends in*
 2036 *Cognitive Sciences*, 10, 278-285.
- 2037 Ghazanfar, A.A. (2009). The multisensory roles for auditory cortex in primate vocal
 2038 communication. *Hearing Research*, 258, 113-120.
- 2039 Grammer, K., & Eibl-Eibesfeldt, I. (1990). The ritualization of laughter. In W. Koch (Ed.),
 2040 *Natürlichkeit der sprache und der kultur: Acta colloquii* (pp. 192-214). Bochum: Brockmeyer.
- 2041 Granot, R., Frankel, Y., Gritsenko, V., Lerer, E., Gritsenko, I., Bachner-Melman, R., S. Israel,
 2042 & Ebstein, R. (2007). Provisional evidence that the arginine vasopressin 1a receptor gene is
 2043 associated with musical memory. *Evolution and Human Behavior*, 28, 313-318.

- 2044 Grape, C., Sandgren, M., Hansson, L., Ericson, M., & Theorell, T. (2003). Does singing
2045 promote well-being?: An empirical study of professional and amateur singers during a singing
2046 lesson. *Integrative Physiological & Behavioral Science*, 38, 65-74.
- 2047 Green, S.R., Mercado, E., Pack, A.A., & Herman, L. (2011). Recurring patterns in the songs
2048 of humpback whales (*Megaptera novaeangliae*). *Behavioural Processes*, 86, 284-294.
- 2049 Greenfield Michael. 2006. Mechanisms and evolution of chorusing interactions in acoustic
2050 insects and anurans. *The Journal of the Acoustical Society of America* 119, 3222.
- 2051 Greenfield, M. D., and Schul, J. (2008). Mechanisms and evolution of synchronous
2052 chorusing: emergent properties and adaptive functions in *Neoconocephalus katydid*s
2053 (Orthoptera: Tettigoniidae). *J. Comp. Physiol. A* 122, 289–297.
- 2054 Griffiths, N., & Reay, J. (2018). The Relative Importance of Aural and Visual Information in
2055 the Evaluation of Western Canon Music Performance by Musicians and Nonmusicians. *Music*
2056 *Perception: An Interdisciplinary Journal*, 35, 364-375.
- 2057 Gruber, T. (2013). Wild-Born Orangutans (*Pongo abelii*) Engage in Triadic Interactions
2058 During Play. *International Journal of Primatology*, 35, 411-424.
- 2059 Hagen, E., & Bryant, G. (2003). Music and dance as a coalition signaling system. *Human*
2060 *Nature*, 14, 21-51.
- 2061 Hagen, E., & Hammerstein, P. (2009). Did Neanderthals and other early humans sing?
2062 Seeking the biological roots of music in the territorial advertisements of primates, lions,
2063 hyenas, and wolves. *Musicae Scientiae*, 13, 291 - 320.
- 2064 Hall, K., Oram, M., Campbell, M.W., Eppley, T., Byrne, R., & Waal, F.D. (2017).
2065 Chimpanzee uses manipulative gaze cues to conceal and reveal information to foraging
2066 competitor. *American Journal of Primatology*, 79, 1-11.
- 2067 Handel, S, Todd, SK and Zoidis, A. 2009. Rhythmic structure in humpback whale
2068 (*Megaptera novaeangliae*) songs: Preliminary implications for song production and

- 2069 perception. *Journal of the Acoustical Society of America, Express Letters*, 125: EL225–
 2070 EL231.
- 2071 Handel, S., Todd, S. K., & Zoidis, A. M. (2012). Hierarchical and rhythmic organization in
 2072 the songs of humpback whales (*Megaptera novaeangliae*). *Bioacoustics*, 21(2), 141–156.
 2073 <https://doi.org/10.1080/09524622.2012.668324>.
- 2074 Hannon, E., & Johnson, S. (2005). Infants use meter to categorize rhythms and melodies:
 2075 Implications for musical structure learning. *Cognitive Psychology*, 50, 354-377.
- 2076 Hannon, E., Vanden Bosch der Nederlanden, C., & Tichko, P. (2012). Effects of perceptual
 2077 experi- ence on children's and adults' perception of unfamiliar rhythms. *Annals of the New*
 2078 *York Academy of Sciences*, 1252, 92–99.
- 2079 Hannon, E., Schachner, A., & Nave-Blodgett, J.E. (2017). Babies know bad dancing when
 2080 they see it: Older but not younger infants discriminate between synchronous and
 2081 asynchronous audiovisual musical displays. *Journal of experimental child psychology*, 159,
 2082 159-174 .
- 2083 Hartbauer, M., & Römer, H. (2016). Rhythm Generation and Rhythm Perception in Insects:
 2084 The Evolution of Synchronous Choruses. *Frontiers in Neuroscience*, 10.
- 2085 Hartbauer, M., Haitzinger, L., Kainz, M., and Römer, H. (2014). Competition and cooperation
 2086 in a synchronous bushcricket chorus. *R. Soc. Open Sci.* 1:140167.
- 2087 Hasegawa, A., Okanoya, K., Hasegawa, T., & Seki, Y. (2011). Rhythmic synchronization
 2088 tapping to an audio–visual metronome in budgerigars. *Scientific Reports*, 1.
- 2089 Hattori, Y., & Tomonaga, M. (2019). Rhythmic swaying induced by sound in chimpanzees
 2090 (Pan troglodytes). *Proceedings of the National Academy of Sciences of the United States of*
 2091 *America*, 117, 936 - 942.
- 2092 Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an
 2093 auditory rhythm in a chimpanzee. *Scientific Reports*, 3.

- 2094 Hay, J., & Diehl, R. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic
 2095 law. *Perception & Psychophysics*, 69, 113-122.
- 2096 Heinsohn, R., Zdenek, C., Cunningham, R., Endler, J., & Langmore, N. (2017). Tool-assisted
 2097 rhythmic drumming in palm cockatoos shares key elements of human instrumental music.
 2098 *Science Advances*, 3.
- 2099 Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., et al.(2010).
 2100 Musical training induces functional plasticity in human hippocampus. *J. Neurosci.* 30, 1377–
 2101 1384.
- 2102 Herzing, D., & Dolphin, W. (2015). Synchronous and Rhythmic Vocalizations and Correlated
 2103 Underwater Behavior of Free-ranging Atlantic Spotted Dolphins (*Stenella frontalis*) and
 2104 Bottlenose Dolphins (*Tursiops truncatus*) in the Bahamas. *Animal Behavior and Cognition*, 2,
 2105 14-29.
- 2106 Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals*
 2107 *of the New York Academy of Sciences*, 1252, 85-91.
- 2108 Honing, H. (2019). The Evolving Animal Orchestra. In search of what makes us musical.
 2109 MIT Press, Cambridge (MA).
- 2110 Honing, H., & Ploeger, A. (2012). Cognition and the Evolution of Music: Pitfalls and
 2111 Prospects. *Topics in cognitive science*, 4 4, 513-24.
- 2112 Honing, H., Cate, C.T., Peretz, I., & Trehub, S. (2015). Without it no music: cognition,
 2113 biology and evolution of musicality. *Philosophical Transactions of the Royal Society B:*
 2114 *Biological Sciences*, 370.
- 2115 Honing, H., Bouwer, F.L., Prado, L., & Merchant, H. (2018). Rhesus Monkeys (*Macaca*
 2116 *mulatta*) Sense Isochrony in Rhythm, but Not the Beat: Additional Support for the Gradual
 2117 Audiomotor Evolution Hypothesis. *Frontiers in Neuroscience*, 12, 475.
- 2118 Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases

- 2119 affiliation. *Social Cognition*, 27(6), 949–961.
- 2120 Hove, M., Marie, C., Bruce, I., & Trainor, L. (2014). Superior time perception for lower
2121 musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings*
2122 *of the National Academy of Sciences*, 111, 10383 - 10388.
- 2123 Hribar, A. (2011). Comparing Great Apes' and Children's Ability to Reason by Analogy. PhD
2124 Thesis, Heidelberg University.
- 2125 Hulse, S.H., Humpal, J., & Cynx, J. (1984). Processing of Rhythmic Sound Structures by
2126 Birds a. *Annals of the New York Academy of Sciences*, 423.
- 2127 Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of*
2128 *Sciences*, 930, 43-61.
- 2129 Huron, D. (2012). Two Challenges in Cognitive Musicology. *Topics in cognitive science*, 4 4,
2130 678-84.
- 2131 Iversen, J., Patel, A., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on
2132 auditory experience. *The Journal of the Acoustical Society of America*, 124 4, 2263-71.
- 2133 Iversen, J., Patel, A., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and
2134 visual rhythms in hearing and deaf individuals. *Cognition*, 134, 232-244.
- 2135 Jacoby, N., & McDermott, J. (2017). Integer Ratio Priors on Musical Rhythm Revealed
2136 Cross-culturally by Iterated Reproduction. *Current Biology*, 27, 359-370.
- 2137 Janata, P. (2009). The Neural Architecture of Music-Evoked Autobiographical
2138 Memories. *Cerebral Cortex (New York, NY)*, 19, 2579 - 2594.
- 2139 Jarvis, E., & Mello, C. (2000). Molecular mapping of brain areas involved in parrot vocal
2140 communication. *Journal of Comparative Neurology*, 419, 1-31.
- 2141 Jespersen, O. (1922). *Language: Its Nature, Development and Origin*. New York, NY:
2142 Norton.
- 2143 Justus, T., & Hutsler, J. J. (2005). Fundamental issues in the evolutionary psychology of

- 2144 music: Assessing innateness and domain specificity. *Music Perception*, 23(1), 1–27.
- 2145 Kello, C., Bella, S.D., Médé, B., & Balasubramaniam, R. (2017). Hierarchical temporal
2146 structure in music, speech and animal vocalizations: jazz is like a conversation, humpbacks
2147 sing like hermit thrushes. *Journal of The Royal Society Interface*, 14, 20170231.
- 2148 Keverne, E.B., Martensz, N., & Tuite, B. (1989). Beta-endorphin concentrations in
2149 cerebrospinal fluid of monkeys are influenced by grooming relationships.
2150 *Psychoneuroendocrinology*, 14, 155-161.
- 2151 Kirschner, S., & Tomasello, M. (2009). Joint drumming: social context facilitates
2152 synchronization in preschool children. *Journal of experimental child psychology*, 102 3, 299-
2153 314.
- 2154 Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-
2155 year-old children. *Evolution and Human Behavior*, 31(5), 354–364.
- 2156 Kisilvesky, B.S., & Muir, D. (1991). Human fetal and subsequent newborn responses to
2157 sound and vibration. *Infant Behavior & Development*, 14, 1-26.
- 2158 Kisilevsky, B. S., Muir, D. W., & Low, J. A. (1992). Maturation of human fetal responses to
2159 vibroacoustic stimulation. *Child Development*, 63(6), 1497–1508.
- 2160 Knörnschild, M., Blüml, S., Steidl, P. *et al.* 2017. Bat songs as acoustic beacons - male
2161 territorial songs attract dispersing females. *Sci Rep* 7, 13918.
- 2162 Koelsch, S. (2009). A Neuroscientific Perspective on Music Therapy. *Annals of the New York*
2163 *Academy of Sciences*, 1169.
- 2164 Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews*
2165 *Neuroscience*, 15, 170-180.
- 2166 Kogan, N. (1997). Reflections on aesthetics and evolution. *Critical Review*, 11, 193-210.
- 2167 Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized Drumming Enhances
2168 Activity in the Caudate and Facilitates Prosocial Commitment - If the Rhythm Comes Easily.

- 2169 *PLoS ONE*, 6.
- 2170 Koshimori, Y., Strafella, A., Valli, M., Sharma, V., Cho, S., Houle, S., & Thaut, M. (2019).
 2171 Motor Synchronization to Rhythmic Auditory Stimulation (RAS) Attenuates Dopaminergic
 2172 Responses in Ventral Striatum in Young Healthy Adults: [11C]-(+)-PHNO PET
 2173 Study. *Frontiers in Neuroscience*, 13.
- 2174 Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing.
 2175 *Trends in Cognitive Sciences*, 22(10), 896–910.
- 2176 Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate
 2177 that other individuals will act according to false beliefs. *Science*, 354, 110 - 114.
- 2178 Lakens, D. (2010). Movement synchrony and perceived entitativity. *Journal of Experimental*
 2179 *Social Psychology*, 46, 701-708.
- 2180 Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: Movement
 2181 synchrony leads to attributions of rapport and entitativity. *Social Cognition*, 29(1), 1–14.
 2182 <https://doi.org/10.1521/soco.2011.29.1.1>.
- 2183 Lameira, A. (2017). Bidding evidence for primate vocal learning and the cultural substrates
 2184 for speech evolution. *Neuroscience & Biobehavioral Reviews*, 83, 429-439.
- 2185 Lameira, A., Eerola, T., & Ravignani, A. (2019). Coupled whole-body rhythmic entrainment
 2186 between two chimpanzees. *Scientific Reports*, 9.
- 2187 Landis-Shack, N., Heinz, A., & Bonn-Miller, M.O. (2017). Music Therapy for Posttraumatic
 2188 Stress in Adults: A Theoretical Review. *Psychomusicology: Music, Mind, and Brain*, 27,
 2189 334–342.
- 2190 Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-
 2191 varying events. *Psychological Review*, 106(1), 119–159.
- 2192 Large, E., & Gray, P.M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo
 2193 (*Pan paniscus*). *Journal of comparative psychology*, 129 4, 317-28.

- 2194 Larsson, M. (2013). Self-generated sounds of locomotion and ventilation and the evolution of
2195 human rhythmic abilities. *Animal Cognition*, 17, 1 - 14.
- 2196 Larsson, M., & Abbott, B. (2018). Is the Capacity for Vocal Learning in Vertebrates Rooted
2197 in Fish Schooling Behavior? *Evolutionary Biology*, 45, 359 - 373.
- 2198 Larsson, M., Richter, J., & Ravignani, A. (2019). Bipedal Steps in the Development of
2199 Rhythmic Behavior in Humans. *Music & Science*, 2.
- 2200 Launay, J., Tarr, B., & Dunbar, R. (2016). Synchrony as an Adaptive Mechanism for Large-
2201 Scale Human Social Bonding. *Ethology*, 122, 779-789.
- 2202 Leanderson, R., Sundberg, J., & Euler, C.V. (1987). Breathing muscle activity and subglottal
2203 pressure dynamics in singing and speech. *Journal of Voice*, 1, 258-261.
- 2204 Lecanuet, J., Granieri-Deferre, C., Jacquet, A., & DeCasper, A. (2000). Fetal discrimination
2205 of low-pitched musical notes. *Developmental psychobiology*, 36 1, 29-39.
- 2206 Lee, N., Elias, D. O., and Mason, A. C. (2009). A precedence effect resolves phantom sound
2207 source illusions in the parasitoid fly *Ormia ochracea*. *Proc. Natl. Acad. Sci. U.S.A.* 106,
2208 6357–6362.
- 2209 Lehmann, J., Korstjens, A. H., & Dunbar, R. I. M. (2007). Group size, grooming and social
2210 cohesion in primates. *Animal Behaviour*, 74(6), 1617–1629.
- 2211 Leman, M., Moelants, D., Varewyck, M., Styns, F., Noorden, L.V., & Martens, J. (2013).
2212 Activating and Relaxing Music Entrain the Speed of Beat Synchronized Walking. *PLoS*
2213 *ONE*, 8.
- 2214 Leubner, D., & Hinterberger, T. (2017). Reviewing the Effectiveness of Music Interventions
2215 in Treating Depression. *Frontiers in Psychology*, 8.
- 2216 Loersch, C., & Arbuckle, N. L. (2013). Unraveling the mystery of music: Music as an evolved
2217 group process. *Journal of Personality and Social Psychology*, 105(5), 777–798.
- 2218 Logue, D.M., & Hall, M. (2014). Migration and the evolution of duetting in songbirds.

- 2219 *Proceedings of the Royal Society B: Biological Sciences*, 281.
- 2220 Lopresti-Goodman, S.M., Kameka, M., & Dube, A. (2013). Stereotypical Behaviors in
- 2221 Chimpanzees Rescued from the African Bushmeat and Pet Trade. *Behavioral Sciences*, 3, 1 -
- 2222 20.
- 2223 Ma, W., & Thompson, W. (2015). Human emotions track changes in the acoustic
- 2224 environment. *Proceedings of the National Academy of Sciences*, 112, 14563 - 14568.
- 2225 Ma, W., Fiveash, A., & Thompson, W. (2019). Spontaneous emergence of language-like and
- 2226 music-like vocalizations from an artificial protolanguage. *Semiotica*, 2019, 1 - 23.
- 2227 MacDougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: The
- 2228 spontaneous tempo of human locomotion. *Journal of Applied Physiology*, 99, 1164–1173.
- 2229 Machin, A., & Dunbar, R. (2011). The brain opioid theory of social attachment: a review of
- 2230 the evidence. *Behaviour*, 148, 985-1025.
- 2231 MacLean, E., & Hare, B.A. (2013). Spontaneous triadic engagement in bonobos (*Pan*
- 2232 *paniscus*) and chimpanzees (*Pan troglodytes*). *Journal of comparative psychology*, 127 3, 245-
- 2233 255.
- 2234 Madison, G. (2014). Sensori-motor synchronisation variability decreases as the number of
- 2235 metrical levels in the stimulus signal increases. *Acta psychologica*, 147, 10-6.
- 2236 Madison, G., Holmquist, J., & Vestin, M. (2018). Musical improvisation skill in a prospective
- 2237 partner is associated with mate value and preferences, consistent with sexual selection and
- 2238 parental investment theory: Implications for the origin of music. *Evolution and Human*
- 2239 *Behavior*, 39(1), 120–129.
- 2240 Malinowski B (1922). *Argonauts of the Western Pacific*. Routledge and Kegan Paul, London.
- 2241 Malinowski, B. (1929). *The sexual life of savages in northwestern Melanesia*. [La vie sexuelle
- 2242 des sauvages du Nord-Ouest de la Mélanésie. (Trans. by S. Jankelevitch.)]. Liveright.
- 2243 Marin, M. M., Schober, R., Gingras, B., & Leder, H. (2017). Misattribution of musical

- 2244 arousal increases sexual attraction towards opposite-sex faces in females. *PLoS ONE*, 12(9),
 2245 Article e0183531.
- 2246 Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals
 2247 Memorize the Rhythm and Timbre of Their Rivals' Voices. *Current Biology*, 27, 2352-
 2248 2356.e2.
- 2249 Matsusaka, T. (2004). When does play panting occur during social play in wild
 2250 chimpanzees? *Primates*, 45, 221-229.
- 2251 McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and
 2252 evolution. *Music Perception*, 23(1), 29–59
- 2253 McPherson, T., Berger, D., Alagapan, S., & Fröhlich, F. (2019). Active and Passive Rhythmic
 2254 Music Therapy Interventions Differentially Modulate Sympathetic Autonomic Nervous
 2255 System Activity. *Journal of music therapy*.
- 2256 Mead M (1976). Growing up in New Guinea. Morrow, New York (original publication 1930).
- 2257 Medrano, L., Salinas, M., Salas, I., Guevara, P., Aguayo, A., Jacobsen, J.K., & Baker, C.
 2258 (1994). Sex identification of humpback whales, *Megaptera novaeangliae*, on the wintering
 2259 grounds of the Mexican Pacific Ocean. *Canadian Journal of Zoology*, 72, 1771-1774.
- 2260 Mehr, S. A., & Krasnow, M. M. (2017). Parent-offspring conflict and the evolution of infant-
 2261 directed song. *Evolution and Human Behavior*, 38(5), 674–684.
- 2262 Mehr, S.A., Singh, M., York, H., Glowacki, L., & Krasnow, M.M. (2018). Form and Function
 2263 in Human Song. *Current Biology*, 28, 356-368.e5.
- 2264 Mehr, S. A., Singh, M., Knox, D., ... Glowacki, L. (2019). Universality and diversity in
 2265 human song. *Science*, 366(6468), 1108 957–970.
- 2266 Mehr, S.A., Krasnow, M.M., Bryant, G., & Hagen, E. (2020). Origins of music in credible
 2267 signaling. *The Behavioral and brain sciences*, 1-41.
- 2268 Menne-Lothmann, C., Jacobs, N., Derom, C., Thiery, E., Os, J., & Wichers, M. (2012).

- 2269 Genetic and Environmental Causes of Individual Differences in Daily Life Positive Affect
 2270 and Reward Experience and Its Overlap with Stress-Sensitivity. *Behavior Genetics*, 42, 778-
 2271 786.
- 2272 Mercado, E. (2018). The Sonar Model for Humpback Whale Song Revised. *Frontiers in*
 2273 *psychology*, 9, 1156.
- 2274 Mercado, Eduardo III, Handel, Stephen. 2012. Understanding the structure of humpback
 2275 whale songs (L). *The Journal of the Acoustical Society of America*, 132:2947-50.
- 2276 Merchant, H., Harrington, D., & Meck, W. (2013). Neural basis of the perception and
 2277 estimation of time. *Annual review of neuroscience*, 36, 313-36.
- 2278 Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic
 2279 entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in*
 2280 *Neuroscience*, 7, 1-8.
- 2281 Merchant, H., Pérez, O., Bartolo, R., Méndez, J., Mendoza, G., Gámez, J., Yc, K., & Prado,
 2282 L. (2015). Sensorimotor neural dynamics during isochronous tapping in the medial premotor
 2283 cortex of the macaque. *European Journal of Neuroscience*, 41, 586-602.
- 2284 Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, &
 2285 S. Brown(Eds.), *The origins of music* (pp. 315–327). Cambridge, Mass: The MIT Press.
- 2286 Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae*, 6, 3–21.
- 2287 Merker, B., Madison, G. S., and Eckerdal, P. (2009). On the role and origin of isochrony in
 2288 human rhythmic entrainment. *Cortex* 45, 4–17. doi: 10.1016/j.cortex.2008.06.011.
- 2289 Merker, B.H., Morley, I., & Zuidema, W.H. (2015). Five fundamental constraints on theories
 2290 of the origins of music. *Philosophical Transactions of the Royal Society B: Biological*
 2291 *Sciences*, 370.
- 2292 Miani A. 2016 a. Sexual arousal and rhythmic synchronization: A possible effect of
 2293 vasopressin. *Med Hypotheses*. 93:122-5.

- 2294 Miani A. 2016 b. The impact of sex and menstrual cycle on rhythmic synchronization an
 2295 experimental approach. MA Thesis, Cognitive Semiotics, Aarhus University.
- 2296 Miller, G. 2000. Evolution of human music through sexual selection. *In* The Origins of
 2297 Music. N.L. Wallin, B. Merker & S.Brown, Eds.: 329-360. MIT Press. Cambridge, MA.
- 2298 Miller, N.S., Kwak, Y., Bohnen, N., Müller, M.L., & Seidler, R. (2013). The pattern of striatal
 2299 dopaminergic denervation explains sensorimotor synchronization accuracy in Parkinson's
 2300 disease. *Behavioural Brain Research*, 257, 100-110.
- 2301 Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: Tests
 2302 of three hypotheses. *Behaviour*, 135(8-9), 1041–1064.
- 2303 Mithen, S. 2005. The singing Neanderthals. The origins of music, language, mind and Body.
 2304 Widenfeld & Nicholson, London.
- 2305 Mora, D.M., Nespore, M., & Toro, J.M. (2013). Do humans and nonhuman animals share the
 2306 grouping principles of the iambic–trochaic law? *Attention, Perception, & Psychophysics*, 75,
 2307 92-100.
- 2308 Morley, I. (2013). The prehistory of music: human evolution, archaeology, and the origins of
 2309 musicality. Oxford: Oxford University Press.
- 2310 Morris, G. K., Kerr, G. E., and Fullard, J. H. (1978). Phonotactic preferences of female
 2311 meadow katydids (Orthoptera: *Tettigoniidae*: *Conocephalus nigropleurum*). *Can. J. Zool.* 56,
 2312 1479–1487. doi: 10.1139/z78-205.
- 2313 Moser, C.J., Lee-Rubin, H., Bainbridge, C., Atwood, S., Simson, J., Knox, D., Glowacki, L.,
 2314 Galbarczyk, A., Jasienska, G., Ross, C., Neff, M.B., Martin, A., Cirelli, L.K., Trehub, S.,
 2315 Song, J., Kim, M., Schachner, A., Vardy, T., Atkinson, Q., Antfolk, J., Madhivanan, P.,
 2316 Siddaiah, A., Placek, C., Salali, G.D., Keestra, S., Singh, M., Collins, S.A., Patton, J., Scaff,
 2317 C., Stieglitz, J., Moya, C., Sagar, R.R., Wood, B.M., Krasnow, M.M., & Mehr, S.A. (2020).
 2318 Acoustic regularities in infant-directed vocalizations across cultures. *bioRxiv*.

- 2319 Mosing, M., Pedersen, N., Madison, G., & Ullén, F. (2014). Genetic Pleiotropy Explains
 2320 Associations between Musical Auditory Discrimination and Intelligence. *PLoS ONE*, 9,
 2321 0113874.
- 2322 Mosing, M., Verweij, K., Madison, G., Pedersen, N.L., Zietsch, B., & Ullén, F. (2015). Did
 2323 sexual selection shape human music? Testing predictions from the sexual selection hypothesis
 2324 of music evolution using a large genetically informative sample of over 10,000
 2325 twins. *Evolution and Human Behavior*, 36, 359-366.
- 2326 Musso, M., Weiller, C., Horn, A., Glauche, V., Umarova, R., Hennig, J., Schneider, A., &
 2327 Rijntjes, M. (2015). A single dual-stream framework for syntactic computations in music and
 2328 language. *NeuroImage*, 117, 267-283.
- 2329 Müller, A., & Anzenberger, G. (2002). Duetting in the Titi Monkey *Callicebus cupreus*:
 2330 Structure, Pair Specificity and Development of Duets. *Folia Primatologica*, 73, 104 - 115.
- 2331 Namu, Yang Erche, Mathieu, Christine. 2004. *Leaving Mother Lake: A Girlhood at the Edge*
 2332 *of the World*, Back Bay Books.
- 2333 Nelson, P.C., & Young, E. (2010). Neural Correlates of Context-Dependent Perceptual
 2334 Enhancement in the Inferior Colliculus. *The Journal of Neuroscience*, 30, 6577 - 6587.
- 2335 Nettl, B. (2015). The study of ethnomusicology: Thirty-three discussions, Urbana, IL:
 2336 University of Illinois Press.
- 2337 Nettl, B. (2000) An Ethnomusicologist Contemplates Universals in Musical Sound and
 2338 Musical Culture. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of music* (pp.
 2339 463–472). Cambridge, Mass: The MIT Press.
- 2340 Niarchou, M., Gustavson, D.E., Sathirapongsasuti, J., Anglada-Tort, M., Eising, E., Bell, E.,
 2341 McArthur, E., Straub, P., McAuley, J., Capra, J.A., Ullén, F., Creanza, N., Mosing, M.,
 2342 Hinds, D., Davis, L., Jacoby, N., & Gordon, R. (2019). Unravelling the genetic architecture of
 2343 musical rhythm: a large-scale genome-wide association study of beat

- 2344 synchronization. *bioRxiv*.
- 2345 Nilsson, U. (2009). Soothing music can increase oxytocin levels during bed rest after open-
- 2346 heart surgery: a randomised control trial. *Journal of clinical nursing*, 18 15, 2153-61.
- 2347 Noda, T., Amemiya, T., Shiramatsu, T.I., & Takahashi, H. (2017). Stimulus Phase Locking of
- 2348 Cortical Oscillations for Rhythmic Tone Sequences in Rats. *Frontiers in Neural Circuits*, 11.
- 2349 O'Bryan, L.R., Munson, B., & Wilson, M. (2018). Chimpanzee food-associated calls display
- 2350 low referential potential in a wild population. *Journal of the Acoustical Society of America*,
- 2351 143, 1788-1788.
- 2352 Obozova, T., Smirnova, A., Zorina, Z., & Wasserman, E. (2015). Analogical reasoning in
- 2353 amazons. *Animal Cognition*, 18, 1363-1371.
- 2354 Oesch, N., & Dunbar, R. (2017). The emergence of recursion in human language: Mentalising
- 2355 predicts recursive syntax task performance. *Journal of Neurolinguistics*, 43, 95-106.
- 2356 Oikkonen, J., Huang, Y., Onkamo, P., Ukkola-Vuoti, L., Raijas, P., Karma, K., Vieland, V.,
- 2357 & Järvelä, I. (2015). A genome-wide linkage and association study of musical aptitude
- 2358 identifies loci containing genes related to inner ear development and neurocognitive
- 2359 functions. *Molecular psychiatry*, 20, 275 - 282.
- 2360 Oikkonen, J., Onkamo, P., Järvelä, I., & Kanduri, C. (2016). Convergent evidence for the
- 2361 molecular basis of musical traits. *Scientific Reports*, 6.
- 2362 Ooishi, Y., Mukai, H., Watanabe, K., Kawato, S., & Kashino, M. (2017). Increase in salivary
- 2363 oxytocin and decrease in salivary cortisol after listening to relaxing slow-tempo and exciting
- 2364 fast-tempo music. *PLoS ONE*, 12.
- 2365 Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: the neuro-affective
- 2366 foundations of musical appreciation. *Behavioural Processes*, 60, 133-155.
- 2367 Parncutt, R., & Chuckrow, R. (2019). Chuckrow's theory of the prenatal origin of music.
- 2368 *Musicae Scientiae*, 23, 403 - 425.

- 2369 Parr, L. A., & Waller, B. M. (2006). Understanding chimpanzee facial expression: Insights
2370 into the evolution of communication. *Social Cognitive and Affective Neuroscience*, 1, 221–
2371 228.
- 2372 Party, Virginie, Brunel-Pons, Odette, Greenfield, Michael D. 2014. Priority of precedence:
2373 receiver psychology, female preference for leading calls and sexual selection in insect
2374 choruses, *Animal Behaviour*, 87: 175-185.
- 2375 Patel, A. (2006). Musical Rhythm, Linguistic Rhythm, and Human Evolution. *Music*
2376 *Perception*, 24, 99-104.
- 2377 Patel, A., Iversen, J., Bregman, M., & Schulz, I. (2009). Experimental Evidence for
2378 Synchronization to a Musical Beat in a Nonhuman Animal. *Current Biology*, 19, 827-830.
- 2379 Pearce, E., Launay, J., MacCarron, P., & Dunbar, R. (2017). Tuning in to others: Exploring
2380 relational and collective bonding in singing and non-singing groups over time. *Psychology of*
2381 *Music*, 45, 496 - 512.
- 2382 Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6,
2383 688-691.
- 2384 Persson, T., Sauciuc, G., & Madsen, E. (2018). Spontaneous cross-species imitation in
2385 interactions between chimpanzees and zoo visitors. *Primates; Journal of Primatology*, 59, 19
2386 - 29.
- 2387 Petkov, C., & Jarvis, E. (2012). Birds, primates, and spoken language origins: behavioral
2388 phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4.
- 2389 Pfenning, A.R., Hara, E., Whitney, O., Rivas, M., Wang, R., Roulhac, P.L., Howard, J.,
2390 Wirthlin, M.E., Lovell, P., Ganapathy, G., Mountcastle, J., Moseley, M., Thompson, J.,
2391 Soderblom, E., Iriki, A., Kato, M., Gilbert, M., Zhang, G., Bakken, T.E., Bongaarts, A.,
2392 Bernard, A., Lein, E.S., Mello, C., Hartemink, A., & Jarvis, E. (2014). Convergent
2393 transcriptional specializations in the brains of humans and song-learning birds. *Science*, 346.

- 2394 Piloizzi, A.R., Carro, C.E., & Huang, X. (2020). Roles of β -Endorphin in Stress, Behavior,
 2395 Neuroinflammation, and Brain Energy Metabolism. *International Journal of Molecular*
 2396 *Sciences*, 22, 338.
- 2397 Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and
 2398 synchronization during the perinatal period. *Frontiers in Psychology*, 5, Article 1048.
- 2399 Pulli, K., Karma, K., Norio, R., Sistonen, P., Göring, H., & Järvelä, I. (2008). Genome-wide
 2400 linkage scan for loci of musical aptitude in Finnish families: evidence for a major locus at
 2401 4q22. *Journal of Medical Genetics*, 45, 451 - 456.
- 2402 Rabinowitch, T.-C., & Meltzoff, A. N. (2017). Synchronized movement experience enhances
 2403 peer cooperation in preschool children. *Journal of Experimental Child Psychology*, 160, 21–
 2404 32. <https://doi.org/10.1016/j.jecp.2017.03.001>.
- 2405 Rabinowitch, T., & Meltzoff, A. (2017). Joint Rhythmic Movement Increases 4-Year-Old
 2406 Children's Prosocial Sharing and Fairness Toward Peers. *Frontiers in Psychology*, 8.
- 2407 Rajendran, V.G., Harper, N.S., & Schnupp, J.W. (2020). Auditory cortical representation of
 2408 music favours the perceived beat. *Royal Society Open Science*, 7.
- 2409 Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human
 2410 rhythm. *Frontiers in Psychology*, 8, Article 1820.
- 2411 Ravignani, A., & Sonnweber, R. (2017). Chimpanzees process structural isomorphisms across
 2412 sensory modalities. *Cognition*, 161, 74 - 79.
- 2413 Ravignani, A., Bowling, D., & Fitch, W. (2014). Chorusing, synchrony, and the evolutionary
 2414 functions of rhythm. *Frontiers in Psychology*, 5.
- 2415 Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits
 2416 rhythmic universals. *Nature Human Behaviour*, 1:0007.
- 2417 Ravignani, A., Chiandetti, C., & Kotz, S. A. (2019). Rhythm and music in animal signals. In
 2418 J. C. Choe (Ed.), *Encyclopedia of animal behavior* (pp. 615–622). Elsevier Academic Press.

- 2419 Ravignani, A., Kello, C., Reus, K.D., Kotz, S., Bella, S.D., Méndez-Aróstegui, M., Rapado-
 2420 Tamarit, B., Rubio-Garcia, A., & Boer, B.D. (2019b). Ontogeny of vocal rhythms in harbor
 2421 seal pups: an exploratory study. *Current Zoology*, 65, 107 - 120.
- 2422 Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized
 2423 prosociality? *Religion, Brain & Behavior*, 4, 19 - 3.
- 2424 Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's Dance Together: Synchrony, Shared
 2425 Intentionality and Cooperation. *PLoS ONE*, 8.Reddy, V. (2015). Joining intentions in
 2426 infancy. *Journal of Consciousness Studies*, 22, 24-44.
- 2427 Repp, B., & Su, Y. (2013). Sensorimotor synchronization: A review of recent research (2006–
 2428 2012). *Psychonomic Bulletin & Review*, 20, 403-452.
- 2429 Richman, B. (1978). The synchronization of voices by gelada monkeys. *Primates*, 19, 569-
 2430 581.
- 2431 Richter, J., and Ostovar, R. (2016). “It don't mean a thing if it ain't got that swing” - an
 2432 alternative concept for understanding the evolution of dance and music inhuman beings.
 2433 *Front. Hum. Neurosci*, 10, art. 485.
- 2434 Riedl, R., & Javor, A. (2012). The Biology of Trust: Integrating Evidence From Genetics,
 2435 Endocrinology, and Functional Brain Imaging. *Journal of Neuroscience, Psychology, and*
 2436 *Economics*, 5, 63-91.
- 2437 Roederer, J. (1984). The Search for a Survival Value of Music. *Music Perception: An*
 2438 *Interdisciplinary Journal*, 1, 350-356.
- 2439 Roeske, T., Tchernichovski, O., Poeppel, D., & Jacoby, N. (2020). Categorical Rhythms Are
 2440 Shared between Songbirds and Humans. *Current Biology*, 30, 3544-3555.e6.
- 2441 Rogers T.L., 2017, 'Calling underwater is a costly signal: Size-related differences in the call
 2442 rates of Antarctic leopard seals', *Current Zoology*, vol. 63, pp. 433 – 443.
- 2443 Rogers, C.N., Ross, A., Sahu, S., Siegel, E., Dooyema, J., Cree, M.A., Stopa, E., Young, L.,

- 2444 Rilling, J., Albers, H., & Preuss, T. (2018). Oxytocin- and arginine vasopressin-containing
 2445 fibers in the cortex of humans, chimpanzees, and rhesus macaques. *American Journal of*
 2446 *Primatology*, 80.
- 2447 Rosenblum, L. (2008). Speech Perception as a Multimodal Phenomenon. *Current Directions*
 2448 *in Psychological Science*, 17, 405 - 409.
- 2449 Rossmanith, N., Costall, A., Reichelt, A. F., López, B., & Reddy, V. (2014). Jointly
 2450 structuring triadic spaces of meaning and action: Book sharing from 3 months on. *Frontiers in*
 2451 *Psychology*, 5, 1390.
- 2452 Rouse, A.A., Cook, P.F., Large, E., & Reichmuth, C. (2016). Beat Keeping in a Sea Lion As
 2453 Coupled Oscillation: Implications for Comparative Understanding of Human Rhythm.
 2454 *Frontiers in Neuroscience*, 10.
- 2455 Rousseau, J.-J. (1781). *Essay on the Origin of Languages*. English translation by J. H. Moran
 2456 and A. Gode (1986). Chicago, IL: University of Chicago Press.
- 2457 Ryan, M., Fox, J., Wilczynski, W. *et al.* Sexual selection for sensory exploitation in the frog
 2458 *Physalaemus pustulosus*. *Nature* 343, 66–67 (1990). <https://doi.org/10.1038/343066a0>
- 2459 Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology*
 2460 *of Music*, 39, 307 - 327.
- 2461 Saarikallio, S., & Erkkilä, J. (2007). The Role of music in adolescents' mood regulation.
 2462 *Psychology of Music*, 35(1), 88–109.
- 2463 Sachs C (1962). The wellsprings of music. M Nijhoff, The Hague.
- 2464 Salimpoor, V., Zald, D., Zatorre, R., Dagher, A., & McIntosh, A. (2015). Predictions and the
 2465 brain: how musical sounds become rewarding. *Trends in Cognitive Sciences*, 19, 86-91.
- 2466 Samuni, L., Preis, A., Mundry, R., Deschner, T., Crockford, C., & Wittig, R. (2016).
 2467 Oxytocin reactivity during intergroup conflict in wild chimpanzees. *Proceedings of the*
 2468 *National Academy of Sciences*, 114, 268 - 273.

- 2469 Sauciuc G.A., Persson T, Madsen E.A. (2017). The social side of imitation in human
 2470 evolution and development. Shared intentionality and imitation games in chimpanzees and 6-
 2471 month old infants. In Arweström Jansson A, Axelsson A, Andreasson R, Billing E. (red.).
 2472 *Proceeding of the 13th SweCog Conference*, Skövde: University of Skövde, pp. 21-23.
- 2473 Sauciuc, G., Klang, K., Zlakowska, J., & Persson, T. (2019a). Social antecedents of rhythmic
 2474 behaviours in chimpanzees (*Pan troglodytes*). *Folia Primatologica*, 91, 290-291.
- 2475 Sauciuc GA, Zlakowska J, Klang KJ, Bååth R, Persson T. (2019b). Social antecedents and
 2476 consequences of rhythmic behaviours in chimpanzees (*Pan troglodytes*), 15th *SweCog*
 2477 Conference, University of Umeå, Sweden, p. 8.
- 2478 Sauciuc, G., Zlakowska, J., Persson, T., Lenninger, S.M., & Madsen, E.A. (2020). Imitation
 2479 recognition and its prosocial effects in 6-month old infants. *PLoS ONE*, 15, 0232717.
- 2480 Savage, P.E., Brown, S., Sakai, E., & Currie, T.E. (2015). Statistical universals reveal the
 2481 structures and functions of human music. *Proceedings of the National Academy of Sciences*,
 2482 112, 8987 - 8992.
- 2483 Savage, P.E., Tierney, A., & Patel, A. (2017). Global Music Recordings Support the Motor
 2484 Constraint Hypothesis for Human and Avian Song Contour. *Music Perception: An*
 2485 *Interdisciplinary Journal*, 34, 327-334.
- 2486 Schachner, A., Brady, T.F., Pepperberg, I., & Hauser, M. (2009). Spontaneous Motor
 2487 Entrainment to Music in Multiple Vocal Mimicking Species. *Current Biology*, 19, 831-836.
- 2488 Schlinger, B., & Brenowitz, E.A. (2002). Neural and Hormonal Control of Birdsong.
 2489 *Hormones, brain and behavior*, 26, 799-839.
- 2490 Schneider, J.N., & Mercado, E. (2019). Characterizing the rhythm and tempo of sound
 2491 production by singing whales. *Bioacoustics*, 28, 239 - 256.
- 2492 Schruth, D.M., Templeton, C.N., & Holman, D. (2019). A definition of song from human
 2493 music universals observed in primate calls. *bioRxiv*.

- 2494 Seeger, A. (1987). *Why Suyá sing: a musical anthropology of an amazonian people*.
 2495 Cambridge: Cambridge University Press.
- 2496 Seesjärvi, E., Särkämö, T., Vuoksima, E., Tervaniemi, M., Peretz, I., & Kaprio, J. (2016).
 2497 The Nature and Nurture of Melody: A Twin Study of Musical Pitch and Rhythm
 2498 Perception. *Behavior Genetics*, 46, 506-515.
- 2499 Selezneva, E., Deike, S., Knyazeva, S., Scheich, H., Brechmann, A., & Brosch, M. (2013).
 2500 Rhythm sensitivity in macaque monkeys. *Frontiers in Systems Neuroscience*, 7, Article 49.
- 2501 Serrano, A., & Terhune, J. (2001). Within-call repetition may be an anti-masking strategy in
 2502 underwater calls of harp seals (*Pagophilus groenlandicus*). *Canadian Journal of Zoology*, 79,
 2503 1410-1413.
- 2504 Sihvonen, A.J., Särkämö, T., Leo, V., Tervaniemi, M., & Soinila, S. (2017). Music-based
 2505 interventions in neurological rehabilitation. *The Lancet Neurology*, 16, 648-660.
- 2506 Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows Spontaneously
 2507 Exhibit Analogical Reasoning. *Current Biology*, 25, 256-260.
- 2508 Sonnweber, R., Ravignani, A., & Fitch, W. (2015). Non-adjacent visual dependency learning
 2509 in chimpanzees. *Animal Cognition*, 18, 733 - 745.
- 2510 Spencer, H. (1857). The origin and function of music. *Fraser's Mag.* 56, 396–408.
- 2511 Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Cambridge, Blackwell.
- 2512 Spierings, M.J., Hubert, J., & Cate, C.T. (2017). Selective auditory grouping by zebra finches:
 2513 testing the iambic–trochaic law. *Animal Cognition*, 20, 665 - 675.
- 2514 Stanwood, G. (2019). Dopamine and Stress. In Fink, G. (ed.), *Stress: Physiology,*
 2515 *Biochemistry, and Pathology*. Pp. 105-114. Academic Press.
- 2516 Stokes, P., Shotbolt, P., Mehta, M., Turkheimer, E., Benecke, A., Copeland, C., Turkheimer,
 2517 F., Lingford-Hughes, A., & Howes, O. (2013). Nature or Nurture? Determining the
 2518 Heritability of Human Striatal Dopamine Function: an [18F]-DOPA PET Study.

- 2519 *Neuropsychopharmacology*, 38, 485-491.
- 2520 Su, Y. (2016). Sensorimotor Synchronization with Different Metrical Levels of Point-Light
- 2521 Dance Movements. *Frontiers in Human Neuroscience*, 10, 186.
- 2522 Sundberg, J. (2018). The Singing voice. In S. Frühholz & P. Berslin (Eds.), *The Oxford*
- 2523 *Handbook of Voice Perception* (pp. 117-142). Oxford University Press.
- 2524 Takeya, R., Kameda, M., Patel, A., & Tanaka, M. (2017). Predictive and tempo-flexible
- 2525 synchronization to a visual metronome in monkeys. *Scientific Reports*, 7.
- 2526 Takeya, R., Patel, A., & Tanaka, M. (2018). Temporal Generalization of Synchronized
- 2527 Saccades Beyond the Trained Range in Monkeys. *Frontiers in Psychology*, 9.
- 2528 Tanner, J., & Byrne, R. (2009). Triadic and collaborative play by gorillas in social games with
- 2529 objects. *Animal Cognition*, 13, 591-607.
- 2530 Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance
- 2531 independently raise pain threshold and encourage social bonding. *Biology Letters*, 11.
- 2532 Tarr, B., Launay, J., & Dunbar, R. (2016). Silent disco: dancing in synchrony leads to
- 2533 elevated pain thresholds and social closeness. *Evolution and human behavior: official journal*
- 2534 *of the Human Behavior and Evolution Society*, 37 5, 343-349.
- 2535 Teki, S., Grube, M., Kumar, S., & Griffiths, T. (2011). Distinct Neural Substrates of
- 2536 Duration-Based and Beat-Based Auditory Timing. *The Journal of Neuroscience*, 31, 3805 -
- 2537 3812.
- 2538 Terleph, T.A., Malaivijitnond, S., & Reichard, U.H. (2017). Male white-handed gibbons
- 2539 flexibly time duet contributions. *Behavioral Ecology and Sociobiology*, 72, 1-13.
- 2540 Thaut, M., McIntosh, G.C., & Hoemberg, V. (2015). Neurobiological foundations of
- 2541 neurologic music therapy: rhythmic entrainment and the motor system. *Frontiers in*
- 2542 *Psychology*, 5.
- 2543 Thayer, R., Newman, J.R., & McClain, T. (1994). Self-regulation of mood: strategies for

- 2544 changing a bad mood, raising energy, and reducing tension. *Journal of personality and social*
 2545 *psychology*, 67 5, 910-25 .
- 2546 Tifferet, S., Gaziel, O., and Baram, Y. (2012). Guitar increases male facebook attractiveness.
 2547 Preliminary support for the sexual selection theory of music. *Lett. Evol. Behav. Sci.* 3, 4–6.
- 2548 Toiviainen, P., Burunat, I., Brattico, E., Vuust, P., & Alluri, V. (2020). The chronnectome of
 2549 musical beat. *NeuroImage*, 216.
- 2550 Tomasello, M., Melis, A.P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two Key Steps
 2551 in the Evolution of Human Cooperation. *Current Anthropology*, 53, 673 - 692.
- 2552 Toro JM (2016) Something old, something new: combining mechanisms during language
 2553 acquisition. *Current Directions in Psychological Science*, 25:130–134.
- 2554 Toro, J.M., & Nespors, M. (2015). Experience-dependent emergence of a grouping bias.
 2555 *Biology Letters*, 11, 20150374.
- 2556 Trainor, L. (2018). The origins of music: auditory scene analysis, evolution, and culture in
 2557 musical creation. In Honing H. (ed.), *The origins of musicality*, MIT, Cambridge MA, pp. 81-
 2558 112.
- 2559 Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Maternal singing in cross-cultural
 2560 perspective. *Infant Behavior and Development*, 16(3), 285–295.
- 2561 Trost, W., Frühholz, S., Schön, D., Labbé, C., Pichon, S., Grandjean, D., & Vuilleumier, P.
 2562 (2014). Getting the beat: Entrainment of brain activity by musical rhythm and pleasantness.
 2563 *NeuroImage*, 103, 55-64.
- 2564 Trost, W., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect
 2565 induction mechanism. *Neuropsychologia*, 96, 96-110.
- 2566 Tsay, C. (2013). Sight over sound in the judgment of music performance. *Proceedings of the*
 2567 *National Academy of Sciences*, 110, 14580 - 14585.
- 2568 Tunçgenç, B., Cohen, E., & Fawcett, C. (2015). Rock with me: the role of movement

- 2569 synchrony in infants' social and nonsocial choices. *Child development*, 86 3, 976-84.
- 2570 Tyack, P. (1981). Interactions between singing Hawaiian humpback whales and conspecifics
2571 nearby. *Behavioral Ecology and Sociobiology*, 8(2), 105-116.
- 2572 Ukkola, L.T., Onkamo, P., Raijas, P., Karma, K., & Järvelä, I. (2009). Musical Aptitude Is
2573 Associated with AVPR1A-Haplotypes. *PLoS ONE*, 4, 0005534.
- 2574 Ullén, F., Mosing, M., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric
2575 properties and heritability of a new online test for musicality, the Swedish Musical
2576 Discrimination Test. *Personality and Individual Differences*, 63, 87-93.
- 2577 Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony
2578 promotes cooperative ability. *Journal of Experimental Social Psychology*, 46(4), 693–695.
- 2579 Valentova, J., Tureček, P., Varella, M.A., Šebesta, P., Mendes, F.D., Pereira, K.J., Kubicová,
2580 L., Stolařová, P., & Havlíček, J. (2019). Vocal Parameters of Speech and Singing Covary and
2581 Are Related to Vocal Attractiveness, Body Measures, and Sociosexuality: A Cross-Cultural
2582 Study. *Frontiers in Psychology*, 10.
- 2583 Vuust, P., & Kringelbach, M. (2010). The Pleasure of Making Sense of Music.
2584 *Interdisciplinary Science Reviews*, 35, 166 - 182.
- 2585 Vuust, P., Ostergaard, L., Pallesen, K.J., Bailey, C., & Roepstorff, A. (2009). Predictive
2586 coding of music – Brain responses to rhythmic incongruity. *Cortex*, 45, 80-92.
- 2587 Watanabe, J. M., & Smuts, B. B. (1999). Explaining religion without explaining it away:
2588 Trust, truth, and the evolution of cooperation in Roy A. Rappaport's "The Obvious Aspects of
2589 Ritual". *American Anthropologist*, 101(1), 98–112.
- 2590 Webb, S.J.N., Schapiro, S.J., Hopkins, W. D. (2019). Polymorphisms in Vasopressin V1a
2591 Receptor Gene (AVPR1A) are Associated with Play, Affiliative, and Exploratory Behavior in
2592 Chimpanzees (Pan troglodytes). The American Society of Primatologists Conference.
- 2593 Wiltermuth, S.S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*,

- 2594 20, 1-5.
- 2595 Winkler, I., Háden, G.P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect
 2596 the beat in music. *Proceedings of the National Academy of Sciences*, 106, 2468 - 2471.
- 2597 Witte, M.D., Spruit, A., Hooren, S.V., Moonen, X., & Stams, G. (2019). Effects of music
 2598 interventions on stress-related outcomes: a systematic review and two meta-analyses. *Health*
 2599 *Psychology Review*, 14, 294 - 324.
- 2600 Yu, L., & Tomonaga, M. (2018). Effect of visual cues in addition to moderate auditory cues
 2601 on temporal coordination: A comparative study in humans and chimpanzees. *bioRxiv*.
- 2602 Yuhi, T., Kyuta, H., Mori, H., Murakami, C., Furuhashi, K., Okuno, M., Takahashi, M., Fuji,
 2603 D., & Higashida, H. (2017). Salivary Oxytocin Concentration Changes during a Group
 2604 Drumming Intervention for Maltreated School Children. *Brain Sciences*, 7, E152.
- 2605 Zahavi, A. (1975). Mate selection-a selection for a handicap. *Journal of theoretical biology*,
 2606 53 1, 205-14.
- 2607 Zarate, J.M. (2013). The neural control of singing. *Frontiers in Human Neuroscience*, 7, 237.
- 2608 Zarco, W., Merchant, H., Prado, L., & Méndez, J. (2009). Subsecond timing in primates:
 2609 comparison of interval production between human subjects and rhesus monkeys. *Journal of*
 2610 *neurophysiology*, 102 6, 3191-202.
- 2611 Zatorre, R., & Baum, S.R. (2012). Musical Melody and Speech Intonation: Singing a
 2612 Different Tune. *PLoS Biology*, 10.
- 2613 Zatorre, R., & Salimpoor, V. (2013). From perception to pleasure: Music and its neural
 2614 substrates. *Proceedings of the National Academy of Sciences*, 110, 10430 - 10437.
- 2615 Zatorre, R. (2015). Musical pleasure and reward: mechanisms and dysfunction. *Annals of the*
 2616 *New York Academy of Sciences*, 1337.
- 2617 Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings*
 2618 *of the National Academy of Sciences*, 107, 5768 - 5773.

- 2619 Zwamborn, E. M. J., & Whitehead, H. (2017). Repeated call sequences and behavioural
2620 context in long-finned pilot whales off Cape Breton, Nova Scotia, Canada. *Bioacoustics*,
2621 26(2), 169–183.
- 2622 Åkerlund, L., & Gramming, P. (1994). Average loudness level, mean fundamental frequency,
2623 and subglottal pressure: comparison between female singers and nonsingers. *Journal of voice:*
2624 *official journal of the Voice Foundation*, 8 3, 263-70.
- 2625 Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of
2626 singing and speaking. *NeuroImage*, 33, 628-635.
- 2627
- 2628

2629