Musicality in human vocal communication: An evolutionary perspective

Juan David Leongómez[1],*, Jan Havlíček[2] and S. Craig Roberts[3],*

[1]*Human Behaviour Lab, Faculty of Psychology, Universidad El Bosque, Bogota, Colombia.*
[2]*Department of Zoology, Charles University, Prague, Czech Republic*
[3]*Faculty of Natural Sciences, University of Stirling, Stirling, UK.*

ORCID IDs: JDL, 0000-0002-0092-6298; JH, 0000-0003-2374-7337; SCR, 0000-0002-9641-6101

* Authors for correspondence (Juan David Leongómez: jleongomez@unbosque.edu.co; S. Craig Roberts: craig.roberts@stir.ac.uk).

Abstract

Studies show that specific vocal modulations, akin to those of infant-directed speech and perhaps music, play a role in communicating intentions and mental states during human social interaction. Based on this, we propose a model for the evolution of musicality –the capacity to process musical information– in relation to human vocal communication. We suggest that a complex social environment, with strong social bonds, promoted the appearance of musicality-related abilities. These social bonds were not limited to those between offspring and mothers or other carers, although these may have been especially influential in view of altriciality of human infants. The model can be further tested in other species by comparing levels of sociality and complexity of vocal communication. By integrating several theories, our model presents a radically different view of musicality, not limited to specifically musical scenarios, but one in which this capacity originally evolved to aid parent-infant communication and bonding, and even today plays a role, not only in music but also in infant-directed speech (IDS), as well as some adult-directed speech (ADS) contexts.

1. Introduction

Musicologists have commonly rejected the idea of music as a universal phenomenon [e.g. 1–3; but see 4], and therefore the study of its origins has not been often addressed within this discipline. In contrast, scientists from disciplines such as biology, psychology, and anthropology have long-standing interest in this idea, often focusing on the purpose of music and its potential evolutionary origin. This essential difference, which perhaps exists because ethnomusicologists usually look at cultural differences and focus on the specificity of individual musical manifestations, while researchers from other disciplines see music as a universal, human phenomenon, has often prevented communication between these complementary views.

While it seems undeniable that all cultures have some individual musical manifestations (i.e. any musical output, from singing and instrument-playing to clapping and dancing) that can be recognised as music [5], ethnomusicologists often highlight the colossal variation in their social contexts and meanings. This makes the scientific study of music as a human universal, and understanding of its origins, extremely challenging.

Scientists have often focused on *music*, which is a behavioural manifestation – the outcome of any potential adaptations – rather than the adaptations themselves. In other words, it may be more pertinent to examine *musicality*, our ability to process musical information, than *music*. Additionally, musicality consists of different separable mechanisms for production and perception that may have evolved independently [6]. Furthermore, because most theories for the origin of music point to an evolutionary connection between music and language, the domain of musicality might be not limited to music but might also play a role in infant-directed speech (IDS) [7] and perhaps even in adult-directed language.

In this paper, we first outline the problems posed by the study of the origins of music, including possible human universals and the relationship between language and music. We then review the evidence regarding the

evolution of music, addressing the challenges posed by the evolutionary study of music and reviewing the key current ideas on this topic. In so doing, we consider, both animal precursors and theories for the evolution of human acoustic communication, as well as their limitations. Finally, we propose a theoretical model for the evolution of musicality, which provides an explanatory and testable framework for explaining newly emerged findings in human acoustic communication. Beyond IDS, recent research has shown that contextual vocal modulation is important (reviewed in this issue by Hughes and Puts). For example, fundamental frequency variability (commonly measured as f_0SD or f_0CV) is modulated and plays an important role in in courtship contexts [e.g. 8–10]. However, this key role of f_0 variability is not apparent in non-courtship scenarios, including authority ranking relationships [e.g. 11, cf. 12], perhaps indicating context-specific variation in the importance of musicality in human communication. Such variation could help us to discover and define the evolutionary processes that have shaped the emergence of human musicality.

2. The puzzling origins of music

There are clear differences in form and complexity between human and non-human forms of vocal communication, which make the understanding of the evolution of human acoustic communication a huge challenge. For example, Falk [13] asked, "why are we the only animals that talk?". Similarly, Brown and Jordania [14] noted that, among more than 4,500 singing species living today, only humans have "the ability to follow precise rhythmic patterns so as to permit group singing, drumming, and dancing", leading them to ask "What explains the unique place of humans among singing species?".

Finding answers to these questions is made challenging by the lack of clear intermediate stages that allow us to create an accurate picture of the evolutionary process towards modern forms of human acoustic communication. There are no other extant hominin species with varying degrees of acoustic communicative complexity, with which we could compare ourselves. And, as the fossil record does not allow to directly study the acoustic communication of extinct species, we are forced to rely on indirect inferences based on archaeological findings or observations of modern animal species (including humans). For instance, even in the case of Neanderthals, *Homo neanderthalensis*, probably the most studied extinct human species, there seem to be no clear conclusions about their level of acoustic communication [15,16].

Darwin himself portrayed music as one of the most mysterious human abilities [17]. It is a phenomenon with no obvious function but seems to be present in all human cultures [18], whose roots can be certainly traced for a few tens of thousands of years, back to the earliest known musical instruments. The earliest known flute – a complex instrument capable of producing differentiated pitches – was made around 40,000 years ago [19,20], However, music does not depend for its existence on the construction of instruments, as we can sing, dance, and use our bodies or other natural objects as drums and so we can infer that musical ability significantly pre-dates this time.

2.1 The question of music universals

The universality of human music is critical to the study of musicality. If musical production is in fact a universal phenomenon, then the idea of it having a purely cultural origin would be hard to defend, suggesting instead some form of biological antecedent.

This is not to say that culture is unimportant. Unquestionably, cultural distinctiveness and variation define individual features, social roles, meanings, and conceptions of music. As Cross [21] points out, within the humanities, including main trends in musicology and ethnomusicology, there is broad consensus in the view of music as a cultural construction. This view is supported precisely by the enormous cultural variation of musical or music-like phenomena in human societies. Moreover, the notion of *music* itself varies significantly between different cultures. For example, Australian Aboriginal songs combine visual, performing and oral arts [22], while the Igbo concept of *nkwa* includes not only actions like singing and playing instruments, but also dancing [23]. In fact, many scholars prefer to use the term *musics* instead of *music*, to account for the uniqueness of these phenomena within each culture [see 2].

This idea of *musics* as cultural expressions lacking any relevant commonalities, only valid within the context of a particular human group [1,2], is essential to understand the limitations that it places on the study of music. If there are no universals in *musics* – no common basic principles that allow measurements and comparisons to

be made – then the scientific study of music as one universal human phenomenon would be irrelevant and perhaps implausible. However, this view is by no means common to all researchers within the humanities [e.g. 5,14,24]; Blacking [5], for example, stated that every society has some cultural manifestations that can be recognised as music, implying that there are, in fact, common features.

In the light of immense cultural variability, it seems difficult to agree about a definition of what music is, and especially how it arose and *what is it for*. However, despite these obstacles, we know that our brain, physiology, and psychology render humans capable of producing and listening to music. In other words, we do at least know that the capacity to process musical information, *musicality*, is universal.

For some decades, scientist from diverse disciplines have presented data that speak to us of a more primal, biological basis of musicality, common to all humans. And, furthermore, science has provided an insight into the cognitive demands of the musical capacity. We *all* share the amazing capacity to produce, perceive and enjoy – or dislike – music, probably soon after we are born [25–27] (or perhaps even before [e.g. 28,29]), and music has a substantial capacity to affect our emotions [30,31].

In fact, many scientists have proposed a variety of music universals. Fritz et al. [32] found that adult Mafa people (from Cameroon and Nigeria) were successful in identifying three basic emotions (happy, sad, scared/fearful) in Western music, at above chance levels. Both Western and Mafa participants also preferred original versions of Western and Mafa music over spectral manipulations of the originals (that affected the sensory dissonance of the music), suggesting that emotionally laden music can be cross-culturally recognised and that its appreciation is universally affected by consonance and dissonance. Trehub [33] proposed the existence of several universal musical features, based on analyses of responses of human infants and adults to original and transposed melodies. These included the perception of contours (i.e. relational pitch and time features of music), scales composed of unequal steps, and a preference for small integer frequency ratios (i.e. consonances, such as the octave (2:1), perfect fifth (3:2), and perfect fourth (4:3)) over large integer ratios (dissonances, such as the tritone (45:32)). In addition, Trehub suggested the universality of a music genre for infants (e.g. lullabies and play songs). In fact, adults can recognise a lullaby as such, even when they are unfamiliar with the musical culture, and can identify with almost absolute precision when a song was sang to an infant [34–37].

Brown and Jordania [14] proposed an extensive list of music universals, categorised in four types: (1) *conserved universals* which apply to all musical utterances and include pseudo-syntactic elements such as music being organised into phrases, the use of relative pitch elements such as the equivalence of octaves (and consequent transposability of melodies), and the use of discreet pitches, as well as factors used for emotive expression, such as register, tempo, and amplitude; (2) *predominant patterns* which apply to all musical styles, including rhythmic features such as the predominance of isometric rhythms, the use of scales divided into seven or less pitches, the use of motives, and use of texts, among others; (3) *common patterns*, which apply to many styles and include, for example, the association of music and dance, and the use of aerophone instruments (wind instruments); and (4) *range universals*, which contains a set of possible options for all musical systems, such as textures (monophony, heterophony, homophony, or polyphony) and type of arrangement (solo or group arrangements). Many of the musical universals proposed by Brown & Jordania were confirmed empirically by Savage and colleagues through an analysis of 304 recordings of diverse traditional musics from around the world [4].

More recently, and by examining a large sample of societies, Mehr and colleagues have produced evidence showing that musical forms can be described by two dimensions (rhythm and melody), that musical behaviour can be described by three components (arousal, religiosity, and formality), and that tonality may be universal [38]. Furthermore, they showed that there are robust cross-cultural associations between form (for example in dance songs or lullabies) and functions in vocal music (e.g., "for dancing", "used to soothe a baby"), and that these functions can be detected by listeners from around the world [35].

2.2   Music-language relationship

Research into universals of musical form and structure suggests that music and its perception is related in complex ways to language, or at least is analogous to it. The deep relationship between language and music in terms of shared neural resources, is supported by evidence presented in a variety of studies [e.g. 39–41], and has become an important area of research and source of debate in recent years.

An increasing number of studies show an important overlap of neural resources involved in the processing of specific music and language tasks [e.g. 42–44]. For instance, strong evidence for shared resources in musical and linguistic syntactic processing has been shown in several studies [39,44]. Moreover, children who suffer from Specific Language Impairment (SLI), which is characterized by deficient processing of linguistic syntax, also show a deficiency of musical syntax processing [45]. There is even evidence suggesting that the human brain does not treat language and music as different kinds of stimuli, at least in early stages of infancy, when it seems to treat language "as a special case of music" [46]. Furthermore, music therapy has been successfully used in speech rehabilitation [47–50].

As in language, music processing involves networks of extensively distributed brain regions. In fact, music might comprise an even vaster network of regions, from both hemispheres, and with an overall asymmetry towards the right hemisphere for pitch processing [51,52]. Hence, the overlap between the activated neural areas for music and language processing that has been found in several neuroimaging studies – especially clear in production tasks that involve singing with lyrics – is not surprising.

Indeed, Peretz [52] points out that in this context – in which overlapping of involved neural resources is expected – finding distinct areas of activation for music and language (particularly singing and speaking) can be more enlightening than describing overlaps. Several studies [53–59; for an example of activation in non-singing musical tasks, see 60] have reported, in addition to the expected overlapping, activation of distinct areas for speech and song production. For example, when singing compared to speaking, there is additional activation on the right superior temporal gyrus and in the primary sensorimotor cortex [59]. Furthermore, evidence of domain – specificity of music and language processing becomes apparent from the study of specific cases of brain damage or developmental disorders [26,52,61,62], in which patients might lose musical abilities while maintaining their speaking capacity, like some amusic patients [26,52,61–63], or when patients can sing or play music, but can no longer speak, as in the case of some aphasias [64,65].

What does this deep relationship tell us about the origins of music and language? Is it possible to think that both channels have common origins? Some genetic evidence seems to suggest that this is precisely the case. For example, Alcock et al. [66] found that the FOXP2 gene – which plays a crucial role in the neural development necessary for language and speech – seems to affect rhythm perception and production, while not affecting pitch perceptual and production skills (which seem to be affected by independent genetic factors, as congenital amusia shows [62]). Furthermore, performance in detecting out–of–key notes in popular melodies showed a stronger correlation between identical ($r = .79$) than fraternal ($r = .46$) twins, suggesting that genetic influence – with a heritability of 70–80% – is more important than shared environments for musical pitch perception [67,68].

The findings regarding similarities and differences found in the processing of music and language have led to an interesting debate. While Peretz [52,69] interprets a variety of data as evidence for more complex and specialised cognitive processing requirements than previously thought, and even modularity, pointing to a biological basis of musicality and some form of natural selection, Patel [70] argues that universality and processing specialisation can be explained without evolutionary adaptation. Patel gives the example of the ability to make fire, which, although an invention, "extends deep into our species' past and is found in every human culture" and "provides things that are universally valued by humans, including the ability to cook food, keep warm, and see in dark places". He also highlights the example of reading and writing – both cultural inventions – which are each partially associated with functional specializations in specific brain regions (product of neural plasticity) and in which disorders may sometimes be driven by genetic causes, at least for reading. However, musicality (unlike making fire, reading, writing, or even music), is not a behaviour *per se* but a capacity that seems not to be taught or learned and appears to be present since early infancy [71–74], or even before delivery [28,29]. Thus, the question of whether music is an adaptation could be a dead end [see 75], but the origin of musicality is anything but.

3.   The evolutionary study of music

Over more than two decades, researchers have focused their attention on the evolution of music, producing a great variety of evolutionary theories that range from Pinker's controversial description of music as "evolutionary cheesecake" [76] to purely adaptationist views [77,78]. Because these ideas have been reviewed and discussed elsewhere [e.g. 6,79,80], we will not examine them in depth. Instead, this section addresses: (1)

issues in the evolutionary study of music, and (2) some major ideas in the evolutionary theories of music, enabling us, in section 4, to propose a theoretical model for the evolution of musicality.

3.1   Difficulties in the evolutionary study of music

Besides the fact that music does not seem to play an obvious direct role of biological relevance for survival and reproduction, the evolutionary study of music must face the problem that musicality is likely to consist of different, relatively independent components. Strong evidence for this can be found in the cases in which a disorder affects either pitch or rhythm processing, but not both [for reviews, see 52,66,81,82], indicating the independence of these modules. This means, as Fitch [6] points out, that different components of musicality might have followed independent evolutionary paths, and that "questions like 'When did music evolve?' or 'What is music for?' seem unlikely to have simple unitary answers".

In addition, Justus & Hutsler [83] suggest that the evolutionary study of music might have been somewhat biased, favouring explanations based on natural selection over those involving cultural transmission. This is because most of the recent studies of the origins of music have been based on the approach of evolutionary psychology. This has required researchers to define criteria to assess whether music emerged as an adaptation (i.e. limited by innate factors, domain-specificity, and conferring survival or reproductive advantages), or as an exaptation [83–85]. The problem, however, is yet more complex, as musicality (being a higher-level cognitive domain such as language) probably involves both exaptations and adaptations, making the limits between them quite vague [83,85].

To eventually obtain a complete picture of the evolution of music, both biological (e.g. cognition, mother-infant interactions) and cultural (e.g. learned aesthetic preferences) aspects should be considered. However, these are so intimately connected in any musical manifestation or its perception that disentangling them is challenging. One possible solution is to study infants, assuming them to be individuals who have very limited cultural experience. They can then be compared to adults to infer what does or does not require previous experience. This approach has provided important answers [71–74], but could intrinsically favour hypotheses related to the evolution of musicality from a parent-infant perspective.

3.2   Key ideas in evolutionary theories of music

Evolutionary theories of music are often linked to those of language. There are at least two main aspects that can be discussed separately: (1) the link between animal precursors and human music and language channels, and (2) the evolution of human acoustic communication (including music and language).

3.2.1   Animal precursors

Non-human acoustic communication has been compared to both language and music. In fact, vocalisations from many species are often called songs, because of their complexity and because they are learned [6]. However, except for gibbons, tarsiers, indri, and perhaps also marmosets, tamarins, and titi monkeys [see e.g., 86,87], these complex song-like vocalisations occur only in birds and non-primate mammals such as cetaceans, suggesting that they do not share a common evolutionary path with music, or any other learned, complex human acoustic signals like language [but see e.g., 88]. There are, however, certain similarities and potential instances of convergent evolution that can provide models for the evolution of human acoustic communication [see 6,89]. Vocal learning, for example, seems to work in an analogous way in songbirds and humans, including the prevalence of local dialects and apparent individual differences [90]. In fact, bird brain areas involved in vocal learning have been compared to Broca and Wernicke regions of the human brain, as they activate when a bird hears or sings a song [91,92].

Furthermore, there are interesting parallels between human music and language with vocal signals of other animal species, particularly in instances where animal vocalisations have semantic- and syntax-like elements [for reviews, see e.g., 93–95]. Interesting cases of semantic-like elements (i.e. calls that have symbolic functions) come from domestic chickens, *Gallus domesticus*. Marler and collaborators, for example, showed that food calls produced by males are dependent on the quality of the food, and that females respond selectively to these calls [96,97]. Furthermore, males are sensitive to the audience, producing significantly fewer calls when a rival male is present, and males are more likely to produce dishonest calls (i.e. in the absence of food) when females are further away than when they are nearby [98,99]. Similar sensitivity to social contexts and audiences has been

shown for alarm calls in red junglefowl, *Gallus gallus* [100], which have different calls for different types of predator [101,102].

Perhaps some of the most interesting cases of semantic-like elements in non-human vocal communication, because of phylogenetic proximity with humans, are instances of vocalisations with some degree of symbolic content in primates. For example, vervet monkeys, *Cercopithecus aethiops*, like chickens, have different calls for different types of predators. The presence of jumping, flying and terrestrial predators (like leopards, eagles and pythons) is communicated through different calls to which individuals respond differently: run into trees, look up, or look down, respectively [103]. These distinct calls are evidence of effective categorization of other species, which individuals progressively develop with age and experience. Calls with symbolic functions have also been documented in other primate species [104–107].

Although seemingly simpler in nature, an interesting phenomenon of recombination of vocal elements has been documented in greater spot-nosed monkeys, *Cercopithecus nictitans*. This example is of particular interest because of the semantic-like properties that call series acquire by the recombination of different alarm calls depending on external events [108,109], thus involving both syntactic- and semantic-like properties.

Syntax-like elements have been widely studied and are usually present in species that produce vocalisations that are categorised as songs. Marler [95] divided syntactic elements into two types: (1) phonological syntax (or phonocoding), which is based on the recombination of individual, small phonetic units lacking meaning (e.g. phonemes in human language) to create sequences (e.g. words); and (2) lexical syntax (or lexicoding), in which sequences are recombined to create strings (e.g. sentences) which have meaning both at the sequence (word) and string (sentence) level. While there are large differences in complexity between human and non-human examples, some animal vocalisations have structures that are like those of human language and particularly music, because of the absence of symbolic meaning.

In birds, some species have individual song repertoires with a complexity that exceed that of non-human primates, and that are based on the recombination of elements [for a review, see 6]. Swamp sparrows, *Melospiza georgiana*, for example, have songs that consist of short individual, independent units, which are recombined into different sequences [110]. In winter wrens, *Troglodytes hiemalis*, each male has a repertoire of around 20 songs that incorporate and transform sequences of other winter wren songs [111,112], in a manner that seems to follow a flexible set of rules [113].

A different example of a potential animal precursor is entrainment, the synchronization to external rhythms, which is a phenomenon central to rhythm processing, musicality, and perhaps speech [for a review, see 114]. Entrainment has been experimentally confirmed in at least one individual from another species, a sulphur-crested cockatoo, *Cacatua galerita eleonora* [115,116], and there is evidence of similar behaviour in other species, mainly parrots [117–119] and, importantly, chimpanzees [120–122].

Amongst mammals, however, the most complex vocal behaviour seems to be that of some cetaceans, particularly humpback whales, *Megaptera novaeangliae*, whose songs have syntactic elements analogous to those of songbirds. They are composed of units (analogous to phonemes), which are combined into phrases (relatively fixed sequences of units), and these into themes, which are a collection of phrases (including repetitions and combinations of phrases), which in turn are mixed to form songs with an average duration of 12 to 15 minutes [123]. Furthermore, humpback whale phrases and songs are constantly changing [124]. Like language and music in human cultures, this creates diversification between populations [e.g. 125,126]. Interestingly, the extent and rate of these changes seems to be motivated by novelty [127], and perhaps other factors. Similar changes in individual and population preferences are seen in songbirds [e.g. 128; for a review, see 129].

3.2.2 *The evolution of human acoustic communication*

Probably the best-known theory for the evolution of music, and one that proposes an adaptive function, is that music plays a role in sexual selection [17,78]. For example, the role that birdsongs play in mate choice seems akin to serenading in human societies. Musical ability may also be attractive in itself [130]. Although Fitch [89] highlights that there are no studies showing a positive relation between musical skills and reproductive success or offspring survival, recent studies have provided the empirical support for a theory of the evolution of music through sexual selection: for example, women prefer composers of more complex music around ovulation

[131,132, cf. 133], there is a general increase in perceived mate value according to musical performance quality [134], and there is evidence of genes associated with both musical perception and production [reviewed in 135].

However, music is in no way limited to mating contexts. This may indicate different evolutionary origins. For example, music seems to play an important role in promoting synchronisation and cooperation, as well as group cohesion and identity [see 136]. Because of these social influences, Brown [77] suggests that music may have co-evolved with collective rituals, which could explain the universal association between music and rituals, as well as the rewarding properties of music from a psychological perspective. According to him, based on the capacity of music to promote social cohesion, and because music is overwhelmingly a social phenomenon, the survival value of music is not apparent at an individual, but only at group, level.

A somewhat similar hypothesis is that language evolved as a form of "vocal grooming", to maintain social bonds in increasingly large groups [137–139]. In fact, when phylogenetically controlled, vocal repertoire size strongly predicts group size as well as grooming time in non-human primates [140]. Social bonding is maintained primarily via grooming in primates, but in increasingly large groups this behaviour, which tends to be a one-to-one activity, is less effective. While this theory is presented in relation to the origins of language, it suggests a stage of communal chorusing, lacking propositional meaning, which replaced grooming. Dunbar's theory of vocal grooming is consistent with archaeological evidence, as well as relationships between social group size and neocortex size [see also 141].

Dunbar's hypothesis resembles that of Darwin [17], who suggested a stage of vocal communication in human evolution more closely related to music (singing/humming) than to spoken language. If this is true, music could be a *fossil* of that hypothetical early stage of vocal communication among hominins, often referred to as musical protolanguage [e.g. 142,143] or music-like protolanguage [e.g. 6]. This idea of a shared *common ancestor* between music and language is probably the most recurrent idea in evolutionary musicology. Similar models covering protolanguage stages that relied on musical or music-like elements have been proposed [e.g. 16], including Brown's musilanguage model [144, see also 145], which suggests the idea of an expression spectrum, in which purely referential meaning (lacking emotional content) lies at one end and purely emotional meaning at the other. The main strength of these models, beyond potentially addressing the origins of both music and language, is that they could explain the complex similarities between these (see section 2.2).

Trehub [73] and Dissanayake [146] have suggested that the primary role of music, and songs in particular, is to aid parent-infant communication. This hypothesis is supported by the apparent universality of lullabies [33, see also 35–37], and their calming effects on infants, commonly used to aid them sleep. This idea is also compatible with the existence of IDS and its prevalence in parent-infant interactions; IDS has characteristic vocal modulation patterns which are detected by infants [37,147] and has important effects on language learning [148], communicating affect [149], and strengthening of mother-infant bonds, which could indicate that IDS is an important component in the development of musicality [73,146, see also 150–152]. Furthermore, this hypothesis is compatible with Dunbar's hypothesis of vocal grooming [137–139], and Falk [7,13] suggested that IDS could be a precursor of the social grooming stage that may have led to language.

In comparison to a sexual selection hypothesis for the origin of musicality, a theory based on parent-infant interactions appears to have important advantages. As pointed out by Fitch [89], it can explain the early development of musical perception abilities, as well as the universal existence and effects of lullabies and IDS, therefore providing hints for a model that could explain, not only music and language, but also IDS.

To summarise, all evolutionary theories about musical capacities share an important component of emotional cohesion, invoking benefits obtained during some form of social interaction, whether in mate choice, social bonding within groups, or parent-infant interaction. While evolutionary theories tend to view musicality as a product of one or other of these selection pressures, it is important to contemplate the possibility that different kinds of benefits for musicality may have been in play at different times during its evolution.

3.2.3 *Alternative views*

The two previous sections outline several lines of evidence, based on which we will propose a new model for the evolution of musicality. But before we turn to that task, we should first point out criticisms of some aspects

of this evidence. First, it is essential to mention that correlations among traits do not imply adaptations, and that this is an important limitation of several theories [for a detailed discussion, see 153].

In addition, it is important to consider the genuine possibility that music-like behaviours in non-human animals may be a be poor analogue of human music, especially when comparing it to behaviours and cognitive processes of distant species (perhaps excluding rhythm as, for example, thrush nightingales, *Luscinia luscinia*, have been shown to have universal rhythm categories that are composed of patterns strikingly similar to those of music [154]). In fact, even the perceptual phenomena in (at least some) non-human animals could have very little in common with music. For example, work with European starlings, *Sturnus vulgaris*, has shown that auditory perception in those songbirds vastly differs from that of humans [155; see also 156] and that, unlike humans, sound pattern recognition is based on absolute spectral shapes rather than pitch [157].

Furthermore, despite the prevalence of arguments in support of a sexual selection hypothesis, some of the supporting evidence has been called into question [for a review, see 158]. As highlighted by Mosing et al. [159], the sexual selection hypothesis makes several specific predictions: (1) skilled musicians should have greater mating success than less musically-skilled individuals; (2) musical ability should indicate genetic quality, and therefore there should be an association between musical ability and other putative traits related to genetic quality; and (3) there must be at least partial overlap in the genetic influences in the associations between the first two predictions. In a sample of over 10,000 twins, Mosing et al. found little support for these predictions. Remarkably, they in fact found that musical ability was negatively associated with measures of mating success.

Likewise, in a recent theory for the evolution of music proposed by Mehr et al. [153], the authors pointed to the various weaknesses of a hypothesis of music as a signal of mate quality. In particular, a hypothesis based on a role of music as a signal of male mate quality – which would have co-evolved with female preferences, similar to other species–, predicts that sexually dimorphic signals in the form of music should be emphasised in men. However, there is little evidence of sex differences in musicality [see 153], and, in fact, women may produce more novel songs than men [160].

Similarly, Dunbar's hypothesis of vocal grooming [137–139,161], which is based on a positive correlation between group size and grooming time, has also been heavily criticised in recent years [162,163]. Grueter et al. [162] argue that the association between group size and grooming time is confounded with substrate: they contend that grooming time should be higher in terrestrial (vs arboreal) species, given higher exposure to ectoparasites. They found that terrestriality predicts grooming time better than group size. While these analyses were disputed by Dunbar and Lehmann [164], newer criticism of the vocal grooming hypothesis was presented by Jaeggi et al. [163]; in their analyses, terrestriality was again found to be an important predictor of grooming time, and no evidence of vocal grooming as a less time-consuming, more efficient form of bonding (as stated in the original form of the vocal grooming hypothesis) was found.

Despite these criticisms, the bulk of evidence suggests that there are at least close parallels between forms of vocal communication in human and (some) non-human species, suggesting at least some form of similarity in the selection pressures that have led to similar solutions. In addition, the evidence suggests that music – and particularly rhythm – plays an important role in interactions within human groups.

4. Towards a model for the evolution of musicality

As pointed above, we believe evolutionary theories related to this subject should not primarily focus on music, but on the ability to process musical information (musicality), understanding the cognitive components and potential modules of musicality, and studying their evolutionary history by tracking their role in several domains. It is important to consider that modularity (i.e. organisation into independent components that interact with other components) does not equal domain specificity [52] and, if music, language, and perhaps IDS have a common evolutionary history, then musicality (or components of it) might not be limited to music processing and could in fact play a role in other domains. For example, new evidence suggests musicality affects the perception and imitation of nuanced pronunciations in languages [41].

Musicality seems to integrate several processing modules, of which for at least two there is a common agreement: pitch and rhythm processing [52]. Importantly, most previous models of music evolution do not treat these two components separately [for a notable exception, see 153]. However, due to their relative

independence they may have had separate evolutionary origins, and have been shaped under independent evolutionary processes [6], as discussed in section 3.1.

4.1   Developing the model

As most theories for the origins of music suggest, music and language could be *descendants* of an earlier, vanished form of vocal communication among ancestral hominin species (section 3.2.2). This could help to explain the relationship between music and language (section 2.2) and, potentially, infant-directed speech (IDS). Among these theories, a model based on the role of musicality in infant-parent communication has particular strengths, as it could further explain the universal features of IDS and lullabies, as well as the musicality of babies, not dependent on previous experience [83,84, see also 85].

The most challenging issue for any model, however, is to explain how human populations changed from a state where musicality, and its components, were non-existent or very modest towards one where they become fully developed. Here, we argue that selection on parent-infant communication could have driven such change, because human infants are so vulnerable that even small improvements in communication and bonding might make the difference between life and death. Indeed, the same argument has been made by researchers in a different sensory modality: Wyatt [165] suggests that secretions from the areolar glands on the breast may be the best place to focus in the ongoing search for a human pheromone, because of its potentially critical importance in facilitating successful suckling when the infant is most vulnerable.

Relative pitch (RP) – an important component of pitch processing and musicality [52] – is the ability to process pitches in relation to each other. Without this ability, individuals with different voice registers would be incapable of recognising (or imitating) a melody as such when sang by, for example, children versus adult women or adult men, who tend to have different registers. RP tends to improve with practice [166,167] and may be better in people who speak tonal languages [168]. Across species, RP seems to be rare, but there is evidence of RP in songbirds like the black-capped chickadee, *Poecile atricapillus* [169] and European starlings, *Sturnus vulgaris* [170], and mammals like ferrets, *Mustela putorius furo* [171]. However, to our knowledge, there are no species that can match human RP abilities. Crucially, evidence shows that human neonates are able to process pitch intervals in a similar manner to adults [172]. Like Trainor et al. [173], we suggest that IDS primarily functions to communicate affective states, as supported by the preference of infants for IDS conveying positive over negative emotions [174,175].

Likewise, rhythm processing seems to be of particular importance in terms of synchronisation and entrainment, even from infancy [e.g. 176,177; for a musical entrainment review, see 178]. Today, rhythm processing is important in IDS, consistent with a model based on the role of musicality in infant-parent communication. However, the influence of music in synchronising behaviours and promoting bonding [e.g. 179–182] is consistent with hypotheses based on music playing a role in promoting group cohesion, as in Dunbar's theory of vocal grooming [137–139,161], or with hypotheses in which music functions as a credible signal for coalitional intent [153,183,184]. It is also manifest in group activities: for example, contemporary armies around the world march to music and common analogous examples include rhymes and chants of sports fans and protesters. Neither is this a solely modern phenomenon: notable traditional examples include Zulu War Chants and the Haka from the Māori people of New Zealand. Music seems to reduce physical exertion [185] and increase pain threshold [181,182], which could also partially explain why music is common when human groups perform repetitive, physically demanding tasks.

In other words, rhythm processing seems to enable mechanisms for social bonding and group identification through synchronised behaviours, while pitch processing enables parent-infant bonding and communication through the inferral of emotional states [174,175], laying the ground for the active communication of those states in any form of protolanguage.

In general, the evidence is consistent with a model for the evolution of musicality based on its role in infant-parent communication. Communication between infants and parents is critical to survival: human children are born relatively underdeveloped, parental care is exceptionally long, and children require strong parental bonds to guarantee care and avoid potentially fatal neglect. In fact, IDS is associated with oxytocin levels and other neuropeptides involved in attachment mechanisms [e.g. 186–188]. Better parent-infant communication, and particularly mother-infant bonding, could facilitate social learning in infants, allowing them to acquire the

necessary skills to survive [189]. Moreover, a developmental tool is necessary for language to evolve, and parent-infant communication is crucial in this aspect [as seen today in IDS, 190].

If parent-infant communication promotes infant survival and development, then selection could have acted on individual variation in musicality to the benefit of those with better ability. Moreover, because musicality seems to be at least partially hereditary [see 67,135], adults with a good level of musicality could produce offspring better equipped to process this information and with the potential of being yet more successful parents, adding a new level to the selection pressure for musicality. In fact, there are primate precursors of guided vocal learning at least in marmoset monkeys [191–193], providing grounds for selection to act upon.

Furthermore, this model could integrate the evidence in support for a sexual selection hypothesis, including the preference for composers of more complex music around ovulation [131,132,134; cf. ,153]. If musicality affects infant survival, then cues of musicality would likely become sexually selected. Choosing a partner with musical abilities would be appealing and relevant if it implies a capacity to bond and empathise with infants and other members of the community, and to produce offspring with these abilities. This could explain the role of vocal modulation – such as an increase in f_0 variability, analogous to that of IDS – during courtship [8,10] and its detection and preference by listeners [8].

Sexual selection would, however, require some display of musicality, which could have been manifested in a music-like protolanguage, and would exploit the capacity of music to coordinate behaviour and promote social bonding. In a society where basic forms of group chorusing (proto-songs?) start to appear in the context of social rhythmic and coordinated behaviours, the interaction between the voice of male adults and women or children would tend to create octaves and fifths [33], provided the perceptual preference for these consonances (which in fact does not seem to be unique to humans [for a review, see 194]). These group activities would start to promote, not only social bonding, but also group identity.

The theoretical splitting between music and language from a common ancestor (music-like protolanguage) might have been a product of the increased relevance that propositional syntax [144] and semantics played in human communication, building an increasing specialisation towards language. Nevertheless, language itself is of little use in the context of interactions with pre-linguistic infants, leaving a domain in which musicality remained essential, being uniquely able to communicate and influence emotional states. For this to occur, however, musicality needs to be present in both infants and adults, allowing the cognitive musical abilities to be employed for other purposes in which it remained influential (e.g. group cohesion and social identity).

Although ancestral primate societies lived in groups, and social cohesion and bonding were likely achieved by other means [similar to other primates; e.g. ,137–139], parent-infant communication became fundamental because of increasing altriciality, and the resulting prolonged parental care and need to acquire more and more complex information via social learning. In short, we are suggesting that musicality is based on pre-existing abilities [see also 136], that musicality itself predates music, and that its primary and original purpose was not music (music being, in this view, an epiphenomenon). This model for the evolution of musicality is summarised in Fig. 1.
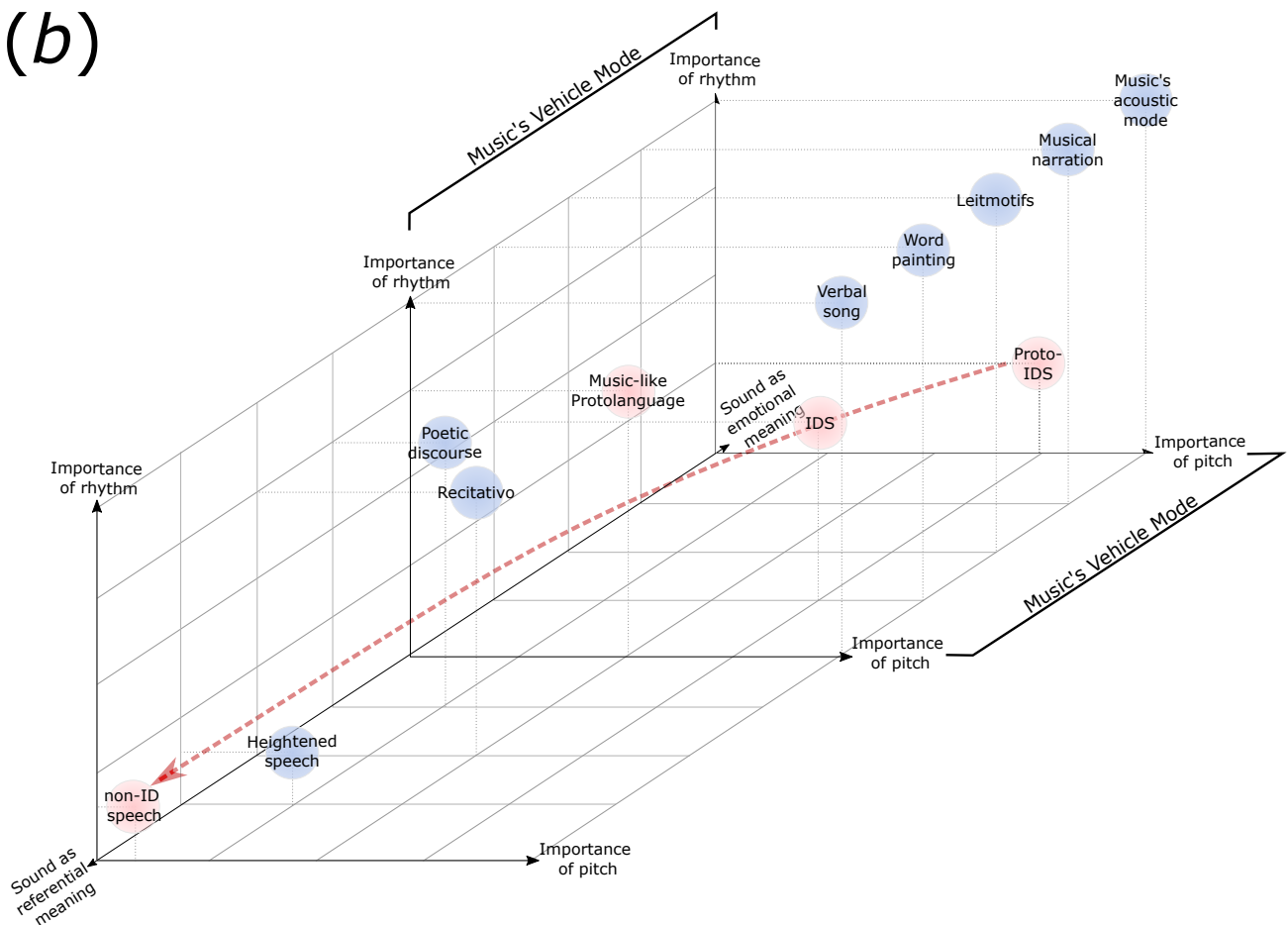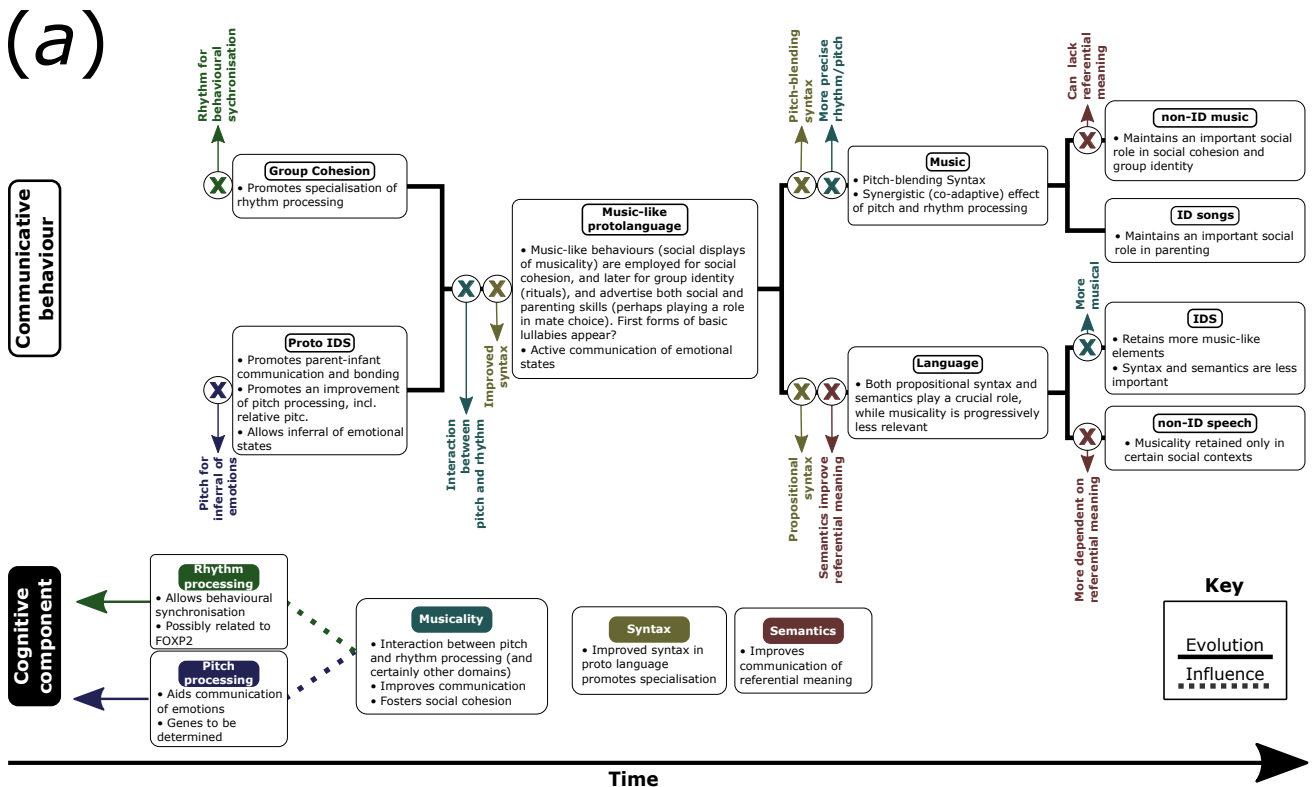
Figure 1. *(a)* Model for the evolution of musicality and its role in human vocal communication. Musicality is presented as a simplified convergence of pitch and rhythm processing, which promotes parent-infant communication and bonding, primarily infant-directed vocalisations (proto-IDS) through pitch, and behavioural synchronisation through rhythm. Music exploits both pitch and rhythm in more precise manners than other forms of communication, and language depends more on propositional syntax and semantics, and less on precise rhythmic or melodic patterns. Coloured Xs represent influences that derive in specialisation and different communicative behaviours. *(b)* Hypothetical location of different functions of human acoustic communication, following Brown's *musilanguage* model [144, Figure 16.1]. In addition to Brown's original

meaning axis (diagonal axis, from referential to emotional), the location of each function is located according to the importance of rhythm (vertical axis) and pitch (horizontal axis) in conveying such meaning. Blue circles represent functions included by Brown, while red circles represent forms included in our model, but not present in the original *musilanguage* model. The red arrow represents the transition in meaning, from emotional to referential, of parent-infant communication becoming more referential as the infant grows. Music forms are represented as having a highly important role of both pith and rhythm, but this can obviously change. For example, in music that promotes group cohesion, like military marches, rhythm tends to have a fundamental role, while pitch's importance can vary widely between melodic marches to exclusively rhythmic ones that lack any melody.

4.2   Implications and tests

If this model accurately portrays the evolution of musicality and its role in complex human vocal communication, musicality could be partially a *fossil* of our musical brain, whose original communicative purpose was bonding and communication between parents and infants, primarily through expressive pitch contours, and later the social communication of intentions and emotions (see Aucouturier in this issue, who shows how vocal cues of expressivity also drive the perception of emotion in music) to promote social bonding and coordination (music-like protolanguage) though an increased importance and precision of rhythm. It is important to emphasise the word *partially*, because its original purpose of aiding communication and bonding between parents and infants is still biologically relevant today in IDS and lullabies, as well as its power to promote group cohesion and social identity, evident today in ritual music, for example.

Moreover, the model incorporates selection pressures that promote the appearance of musicality related abilities, including being able to navigate a complex social environment and develop strong social bonds, particularly those between mother and offspring (which also requires offspring to be born relatively underdeveloped and dependent on their mothers or other adults to survive, incidentally common among many mammal and bird species). These pressures can be tested in other species according to the complexity of their vocal communication. In fact, there is evidence suggesting a positive association between social and vocal complexity in several species including Carolina chickadees, *Poecile carolinensis* [195], sciurid rodents (marmots, *Marmota spp.*, prairie dogs, *Cynomys spp.*, and ground squirrels, *Spermophilus spp.*) [196] as well as other ground-dwelling mammals [197] and bats of the suborder *Microchiroptera* [198]. However, testing species more closely related to humans (i.e. primates), might allow more specific predictions to be tested, regarding the evolutionary pressures that promote vocal complexity [like coevolution of social and vocal complexity, 140], but also those that forged human musicality.

As such, this model predicts that primate species living in particularly complex social environments are the most likely candidates to show musicality-like skills. Indeed, social complexity could be positively associated with more complex communicative systems [199,200], including both vocal diversity and flexibility [e.g. 201,202; see also 153]. For example, bonobo (*Pan paniscus*) vocal interactions respect temporal rules like turn-taking, akin to those of human conversations [88]. Regarding infant-parent bonding, pygmy marmosets, *Cebuella pygmaea*, a socially complex species with cooperative breeding, have been shown to use a form of 'babbling' during infancy that decreases as they grow older [203], and such babbling in both infants as well as juveniles is associated with increased social interactions with other group members [204]. In fact, cooperative breeding has been argued to promote the evolution of language [205,206; see also 86]. However, perhaps the most interesting example comes from gelada baboons, *Theropithecus gelada*, a species that lives in a less complex ecological environment, but a more complex social environment, than the closely related chacma baboons, *Papio ursinus*, and has more vocal complexity [207]. Geladas, in fact, have been shown to synchronise their calls [208], and use both rhythm and *melody* in their vocal interactions [209].

In human language contexts, musicality seems to be required to perceive f_o contours, of which variability in fundamental frequency (f_o variability) modulation seems to be a good indicator [8–10,210]. While most evidence points toward an important role of f_o modulation in courtship contexts and mate choice, new lines of evidence have highlighted the importance of f_o variability modulation and other forms of musicality in more general contexts, including the recently reported association between f_o variability and cooperativeness [211], and the role of musicality in allowing individuals to capture the nuances of linguistic non-verbal aspects [e.g. 41,212].

Recently, two noteworthy but opposing theories for the evolution of music have been proposed. Savage et al.'s [141] model has social bonding as its main function. In contrast, Mehr et al. [153] proposed that music functions as a credible signal for both coalitional interactions (i.e. signalling coalition strength, size, and coordination

ability) and parental attentiveness, functions that evolved from territorial advertisements and contact calls, respectively. In addition, they provide important criticisms of the dominant views of music evolution – that is, that music evolved through sexual selection, to promote social bonding, or as a by-product of other adaptations.

Our model shares some features with both these ideas, but also has some key differences. We fully share Savage et al.'s view of the importance of social bonding, but we believe that music, and especially musicality, need to be placed in a field where communication is much wider than simply music. Furthermore, in contrast to Savage and colleagues, we differentiate the evolution of pitch- and rhythm-related cognition, while their model better fits rhythm only. We also agree with Mehr and colleagues that sexual selection is unlikely to be the primary driver of human musicality, that melody evolved in the context of parent-offspring communication and that rhythm evolved in the context of group interactions. However, if musicality is important for creating and maintaining bonds with infants, or as a signal of parental attention and care [153], cues of this ability may exist in species with alloparenting and/or biparental care, and hence be preferred in potential long-term partners. In humans, this would likely be in the form of voice modulations akin to those of IDS. This would imply that those cues are present in both males and females, as opposed to preferably in males, as a traditional mate quality hypothesis would imply. Thus, our model makes an array of unique predictions, that can be bold, broad, or specific:

(i) Musicality should be associated with better parental abilities and/or disposition, rather than increased mating success.

(ii) Given its importance in several different social contexts, in both infancy and adulthood, we may not observe any large pubertal change in the importance of musicality, despite us normally seeing such changes in other sexually selected traits.

(iii) Musicality itself should not be sexual dimorphic, but sex differences in musical performance, particularly in mating relevant contexts, could exist. It is important to consider that not all sexually selected traits must be sexually dimorphic, particularly in species with mutual mate choice (like preferences for some personality traits such as agreeableness [e.g., 213]).

(iv) In light of alloparenting in humans, because parents are rarely the only caretakers of infants, cues of musicality from potential caretakers may be important as a sign of trust and ability to bond with, and care for, infants.

(v) A core characteristic of IDS (and in fact, "infant-directedness"), as a specific but universal form of vocal modulation, is increased pitch variability [e.g. 37]. Thus, infants should prefer samples of IDS with experimentally increased pitch variability over unmanipulated ones.

(vi) In courtship contexts, pitch variability should correlate with the perceived attractiveness of the target listener, and should be a stronger predictor than mean voice pitch (which has been argued to be a cue of underlying mate quality [e.g. 214]). Preliminary evidence already supports this prediction [10], but it should be especially evident in prospective committed relationships.

(vii) If musicality is cued in non-musical contexts by changes in pitch variability, as indicated in courtship studies of voice modulation [e.g. 8–10], then the ability to detect these subtle modulations should be dependent on pitch discrimination skills; any potential benefit of such vocal modulations would be limited by the listener's capacity to perceive such changes.

(viii) In couples with children, pitch discrimination skills of each parent should be positively associated with satisfaction with their partner as a mother or father.

(ix) Musical rhythm processing should evolve only in species with need for coordination (e.g. cooperation).

13

(x)   Musical pitch processing should evolve only in species with altriciality and long-parental care, where infants' survival chances are linked to the strength of their bonds and ability to communicate with their parents and/or other caretakers.

(xi) Musicality (and perhaps some musical manifestations) should evolve mainly in species with the requirements for predictions ix and x.

The model proposed here is relatively simple, as it is based only on two potential cognitive modules (i.e. pitch and rhythm processing, corresponding to the two principal components found by Mehr et al. [38]), but it provides a general view consistent with the most current evidence and could explain the range of human complex vocal communication. It shares important aspects with previous models, including the idea of a common ancestor to both music and language. However, it also incorporates several novel aspects; in particular, (1) that models' focus should be on musicality rather than music; (2) that modules of musicality, like pitch and rhythm processing, may have different evolutionary paths and could have been shaped by different selection pressures; (3) that musicality may predate music itself; and (4) that musicality could play a role today, not only in music, but also in IDS, and at least certain language contexts like courtship. In addition, we believe that infant-parent bonding is the strongest candidate to explain the emergence of musicality in the human lineage. It might be more than a link between music and language: it could be the very purpose of musicality.

Additional Information

Authors' Contributions J.D.L. proposed the article concept, designed the figures, and drafted the manuscript. J.H. and S.C.R. revised the manuscript critically. All authors developed the concepts, approved the content, and contributed to the writing and ideas.

Competing Interests
The authors have no competing interests.

References

1. Bohlman S. 1999 Ontologies of music. In *Rethinking music* (eds N Cook, M Everist), pp. 17–34. Oxford: OUP.

2. Cross I. 2003 Music and biocultural evolution. In *The Cultural Study of Music: A Critical Introduction* (eds T Herbert, R Middleton), pp. 19–30. New York: Routledge.

3. Tomlinson G. 1984 The web of culture: A context for musicology. *19th-Century Music* **7**, 350–362. (doi:10.2307/746387)

4. Savage PE, Brown S, Sakai E, Currie TE. 2015 Statistical universals reveal the structures and functions of human music. *Proc. Natl. Acad. Sci.* **112**, 8987–8992. (doi:10.1073/pnas.1414495112)

5. Blacking J. 1995 Music, culture and experience. In *Music, culture and experience: selected papers of John Blacking* (ed R Byron), pp. 223–242. London: University of Chicago Press.

6. Fitch WT. 2006 The biology and evolution of music: a comparative perspective. *Cognition* **100**, 173–215. (doi:10.1016/j.cognition.2005.11.009)

7. Falk D. 2004 Prelinguistic evolution in hominin mothers and babies: For cryin'out loud! *Behav. Brain Sci.* **27**, 461–462.

8. Leongómez JD, Binter J, Kubicová L, Stolařová P, Klapilová K, Havlíček J, Roberts SC. 2014 Vocal modulation during courtship increases proceptivity even in naive listeners. *Evol. Hum. Behav.* **35**, 489–496. (doi:10.1016/j.evolhumbehav.2014.06.008)

9. Pisanski K, Oleszkiewicz A, Plachetka J, Gmiterek M, Reby D. 2018 Voice pitch modulation in human mate choice. *Proc. R. Soc. B Biol. Sci.* **285**, 20181634. (doi:10.1098/rspb.2018.1634)

10. Leongómez JD, Sánchez OR, Vásquez-Amézquita M, Roberts SC. 2020 Contextualising courtship: Exploring male body odour effects on vocal modulation. *PsyArXiv* (doi:10.31234/osf.io/8f7g3)

11. Leongómez JD, Mileva VR, Little AC, Roberts SC. 2017 Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics. *PLoS One* **12**, e0179407. (doi:10.1371/journal.pone.0179407)

12. Hodges-Simeon CR, Gaulin SJC, Puts DA. 2010 Different vocal parameters predict perceptions of dominance and attractiveness. *Hum. Nat.* **21**, 406–427. (doi:10.1007/s12110-010-9101-5)

13. Falk D. 2005 Prelinguistic evolution in early hominins: Whence motherese? *Behav. Brain Sci.* **27**, 491–503. (doi:10.1017/S0140525X04000111)

14. Brown S, Jordania J. 2011 Universals in the world's musics. *Psychol. Music* **41**, 229–248. (doi:10.1177/0305735611425896)

15. Dediu D, Levinson SC. 2018 Neanderthal language revisited: not only us. *Curr. Opin. Behav. Sci.* **21**, 49–55. (doi:10.1016/j.cobeha.2018.01.001)

16. Mithen SJ. 2006 *The singing Neanderthals : the origin of music, language, mind and body*. London: Phoenix.

17. Darwin C. 1871 *The descent of man, and selection in relation to sex*. London: Murray.

18. Higgins KM. 2012 *The Music Between Us: Is Music a Universal Language?* Chicago, IL: The University of Chicago Press.

19. Conard NJ, Malina M, Münzel SC. 2009 New flutes document the earliest musical tradition in southwestern Germany. *Nature* **460**, 737–740. (doi:10.1038/nature08169)

20. Adler DS. 2009 Archaeology: The earliest musical tradition. *Nature* **460**, 695–696. (doi:10.1038/460695a)

21. Cross I. 2006 Music, Cognition, Culture, and Evolution. *Ann. N. Y. Acad. Sci.* **930**, 28–42. (doi:10.1111/j.1749-6632.2001.tb05723.x)

22. Ellis CJ. 1984 The Nature of Australian Aboriginal Music. *Int. J. Music Educ.* **4**, 47–50. (doi:10.1177/025576148400400110)

23. Gourlay KA. 1984 The Non-Universality of Music and the Universality of Non-Music. *World Music* **26**, 25–39.

24. Nettl B. 2000 An ethnomusicologist contemplates universals in musical sound and musical culture. In *The Origins of Music* (eds NL Wallin, B Merker, S Brown), pp. 463–472. Cambridge, MA.: MIT Press.

25. Papoušek M. 1996 Intuitive parenting: a hidden source of musical stimulation in infancy. In *Musical Beginnings* (eds I Deliège, J Sloboda), pp. 88–112. Oxford: OUP.

26. Peretz I, Hyde KL. 2003 What is specific to music processing? Insights from congenital amusia. *Trends Cogn. Sci.* **7**, 362–367. (doi:10.1016/S1364-6613(03)00150-5)

27. Perani D, Saccuman MCC, Scifo P, Spada D, Andreolli G, Rovelli R, Baldoli C, Koelsch S. 2010 Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci.* **107**, 4758. (doi:10.1073/pnas.0909074107)

28. Al-Qahtani NH. 2005 Foetal response to music and voice. *Aust. N. Z. J. Obstet. Gynaecol.* **45**, 414–417. (doi:10.1111/j.1479-828x.2005.00458.x)

29. Prochnow A, Erlandsson S, Hesse V, Wermke K. 2019 Does a 'musical' mother tongue influence cry melodies? A comparative study of Swedish and German newborns. *Music. Sci.* **23**, 143–156. (doi:10.1177/1029864917733035)

30. Husain G, Thompson WF, Schellenberg EG. 2002 Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Percept.* **20**, 151–171. (doi:10.1525/mp.2002.20.2.151)

31. Juslin PN, Sloboda JA, editors. 2001 *Music and emotion: Theory and research*. Oxford: Oxford University Press.

32. Fritz T, Jentschke S, Gosselin N, Sammler D, Peretz I, Turner R, Friederici AD, Koelsch S. 2009 Universal recognition of three basic emotions in music. *Curr. Biol.* **19**, 573–576. (doi:10.1016/j.cub.2009.02.058)

33. Trehub SE. 2000 Human processing predispositions and musical universals. In *The Origins of Music* (eds NL Wallin, B Merker, S Brown), pp. 427–448. Cambridge, MA.: MIT Press.

34. Trehub SE, Unyk AM, Trainor LJ. 1993 Maternal singing in cross-cultural perspective. *Infant Behav. Dev.* **16**, 285–295. (doi:10.1016/0163-6383(93)80036-8)

35. Mehr SA, Singh M, York H, Glowacki L, Krasnow MM. 2018 Form and Function in Human Song. *Curr. Biol.* **28**, 356-368.e5. (doi:10.1016/j.cub.2017.12.042)

36. Bainbridge CM *et al.* 2020 Infants relax in response to unfamiliar foreign lullabies. *Nat. Hum. Behav.* (doi:10.1038/s41562-020-00963-z)

37. Moser CJ *et al.* 2020 Acoustic regularities in infant-directed vocalizations across cultures. *bioRxiv* , 2020.04.09.032995. (doi:10.1101/2020.04.09.032995)

38. Mehr SA *et al.* 2019 Universality and diversity in human song. *Science* **366**. (doi:10.1126/science.aax0868)

39. Koelsch S, Fritz T, Schulze K, Alsop D, Schlaug G. 2005 Adults and children processing music: an fMRI study. *NeuroImage* **25**, 1068–1076. (doi:10.1016/j.neuroimage.2004.12.050)

40. Schön D, Magne C, Besson M. 2004 The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* **41**, 341–349.

41. Coumel M, Christiner M, Reiterer SM. 2019 Second Language Accent Faking Ability Depends on Musical Abilities, Not on Working Memory. *Front. Psychol.* **10**, 1–10. (doi:10.3389/fpsyg.2019.00257)

42. Fedorenko E, Patel AD, Casasanto D, Winawer J, Gibson E. 2009 Structural integration in language and music: evidence for a shared system. *Mem. Cognit.* **37**, 1–9. (doi:10.3758/MC.37.1.1)

43. Koelsch S, Gunter TC, Wittfoth M, Sammler D. 2005 Interaction between syntax processing in language and in music: An ERP study. *J. Cogn. Neurosci.* **17**, 1565–1577.

44. Sammler D *et al.* 2009 Overlap of musical and linguistic syntax processing: intracranial ERP evidence. *Ann. N. Y. Acad. Sci.* **1169**, 494–498. (doi:10.1111/j.1749-6632.2009.04792.x)

45. Jentschke S, Koelsch S, Sallat S, Friederici AD. 2008 Children with specific language impairment also show impairment of music-syntactic processing. *J. Cogn. Neurosci.* **20**, 1940–1951. (doi:10.1162/jocn.2008.20135)

46. Koelsch S, Siebel WA. 2005 Towards a neural basis of music perception. *Trends Cogn. Sci.* **9**, 578–84. (doi:10.1016/j.tics.2005.10.001)

47. Skeie GO, Einbu T, Aarli J. 2010 Singing Improves Word Production in Patients with Aphasia. In *Neurology of Music* (ed FC Rose), pp. 347–357. London: Imperial College Press.

48. Hurkmans J, de Bruijn M, Boonstra AM, Jonkers R, Bastiaanse R, Arendzen H, Reinders-Messelink HA. 2012 Music in the treatment of neurological language and speech disorders: A systematic review. *Aphasiology* **26**, 1–19. (doi:10.1080/02687038.2011.602514)

49. Stahl B, Kotz SA, Henseler I, Turner R, Geyer S. 2011 Rhythm in disguise: why singing may not hold the key to recovery from aphasia. *Brain* **134**, 3083–3093. (doi:10.1093/brain/awr240)

50. Racette A, Bard C, Peretz I. 2006 Making non-fluent aphasics speak: sing along! *Brain* **129**, 2571–2584. (doi:10.1093/brain/awl250)

51. Peretz I, Zatorre RJ. 2005 Brain organization for music processing. *Annu. Rev. Psychol.* **56**, 89–114. (doi:10.1146/annurev.psych.56.091103.070225)

52. Peretz I. 2009 Music, language and modularity framed in action. *Psychol. Belg.* **49** **2**, 157–175.

53. Brown S, Martinez MJ, Parsons LM. 2006 Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur. J. Neurosci.* **23**, 2791–2803. (doi:10.1111/j.1460-9568.2006.04785.x)

54. Callan DE, Tsytsarev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H, Turner R. 2006 Song and speech: brain regions involved with perception and covert production. *NeuroImage* **31**, 1327–1342. (doi:10.1016/j.neuroimage.2006.01.036)

55. Hickok G, Buchsbaum B, Humphries C, Muftuler T. 2003 Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* **15**, 673–682. (doi:10.1162/089892903322307393)

56. Jeffries KJ, Fritz JB, Braun AR. 2003 Words in melody: an H215O PET study of brain activation during singing and speaking. *NeuroReport* **14**, 749–754. (doi:10.1097/00001756-200304150-00018)

57. Ozdemir E, Norton A, Schlaug G. 2006 Shared and distinct neural correlates of singing and speaking. *NeuroImage* **33**, 628–635. (doi:10.1016/j.neuroimage.2006.07.013)

58. Saito Y, Ishii K, Yagi K, Tatsumi IF, Mizusawa H. 2006 Cerebral networks for spontaneous and synchronized singing and speaking. *NeuroReport* **17**, 1893–1897.

59. Özdemir E, Norton A, Schlaug G. 2006 Shared and distinct neural correlates of singing and speaking. *NeuroImage* **33**, 628–635. (doi:10.1016/j.neuroimage.2006.07.013)

60. Leonard MK, Desai M, Hungate D, Cai R, Singhal NS, Knowlton RC, Chang EF. 2019 Direct cortical stimulation of inferior frontal cortex disrupts both speech and music production in highly trained musicians. *Cogn. Neuropsychol.* **36**, 158–166. (doi:10.1080/02643294.2018.1472559)

61. Sacks OW. 2007 *Musicophilia: Tales of music and the brain*. London: Picador.

62. Hyde KL, Peretz I. 2004 Brains that are out of tune but in time. *Psychol. Sci.* **15**, 356–360. (doi:10.1111/j.0956-7976.2004.00683.x)

63. Pearce JMS. 2005 Selected observations on amusia. *Eur. Neurol.* **54**, 145–148. (doi:10.1159/000089606)

64. Signoret JL, van Eeckhout P, Poncet M, Castaigne P. 1987 Aphasia without amusia in a blind organist. Verbal alexia-agraphia without musical alexia-agraphia in braille. *Rev. Neurol. (Paris)* **143**, 172–181.

65. Yamadori A, Osumi Y, Masuhara S, Okubo M. 1977 Preservation of singing in Broca's aphasia. *J. Neurol. Neurosurg. Psychiatry* **40**, 221–224. (doi:10.1136/jnnp.40.3.221)

66. Alcock KJ, Passingham RE, Watkins K, Vargha-Khadem F. 2000 Pitch and timing abilities in inherited speech and language impairment. *Brain Lang.* **75**, 34–46. (doi:10.1006/brln.2000.2323)

67. Drayna D, Manichaikul A, De Lange M, Snieder H, Spector T. 2001 Genetic correlates of musical pitch recognition in humans. *Science* **291**, 1969–1972. (doi:10.1126/science.291.5510.1969)

68. Seesjärvi E, Särkämö T, Vuoksimaa E, Tervaniemi M, Peretz I, Kaprio J. 2016 The Nature and Nurture of Melody: A Twin Study of Musical Pitch and Rhythm Perception. *Behav. Genet.* **46**, 506–515. (doi:10.1007/s10519-015-9774-y)

69. Peretz I. 2006 The nature of music from a biological perspective. *Cognition* **100**, 1–32. (doi:10.1016/j.cognition.2005.11.004)

70. Patel AD. 2010 Music, biological evolution, and the brain. In *Emerging Disciplines: Shaping New Fields of Scholarly Inquiry in and beyond the Humanities* (ed M Bailar), pp. 41–64. Houston, TX.: Rice University Press.

71. Bencivelli S. 2011 *Why We Like Music: Ear, Emotion, Evolution*. Hudson, NY: Music Word Media.

72. Trehub SE, Hannon EE. 2006 Infant music perception: domain-general or domain-specific mechanisms? *Cognition* **100**, 73–99. (doi:10.1016/j.cognition.2005.11.006)

73. Trehub SE. 2003 The developmental origins of musicality. *Nat. Neurosci.* **6**, 669–673. (doi:10.1038/nn1084)

74. Trehub SE. 2001 Musical predispositions in infancy. *Ann. N. Y. Acad. Sci.* **930**, 1–16. (doi:10.1111/j.1749-6632.2001.tb05721.x)

75. Fitch WT. 2006 On the biology and evolution of music. *Music Percept.* **24**, 85–88. (doi:10.1525/mp.2006.24.1.85)

76. Pinker S. 1997 *How the mind works*. New York, NY: Norton.

77. Brown S. 2000 Evolutionary models of music: From sexual selection to group selection. In *Perspectives in Ethology* (eds F Tonneau, NS Thompson), pp. 231–281. Kluwer Academic/Plenum.

78. Miller GF. 2000 Evolution of human music through sexual selection. In *The Origins of Music* (eds NL Wallin, B Merker, S Brown), pp. 329–360. MIT Press.

79. Cross I, Morley I. 2008 The evolution of music: Theories, definitions and the nature of the evidence. In *Communicative musicality* (eds SN Malloch, C Trevarthen), pp. 61–82. Oxford: Oxford university Press.

80. Perlovsky L. 2010 Musical emotions: functions, origins, evolution. *Phys. Life Rev.* **7**, 2–27. (doi:10.1016/j.plrev.2009.11.001)

81. Alcock KJ, Wade D, Anslow P, Passingham RE. 2000 Pitch and timing abilities in adult left-hemisphere-dysphasic and right-hemisphere-damaged subjects. *Brain Lang.* **75**, 47–65. (doi:10.1006/brln.2000.2324)

82. Di Pietro M, Laganaro M, Leemann B, Schnider A. 2004 Receptive amusia: temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia* **42**, 868–877. (doi:10.1016/j.neuropsychologia.2003.12.004)

83. Justus T, Hutsler JJ. 2005 Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain specificity. *Music Percept.* **23**, 1–27. (doi:10.1525/mp.2005.23.1.1)

84. McDermott JH, Hauser MD. 2005 The origins of music: Innateness, uniqueness, and evolution. *Music Percept.* **23**, 29–60.

85. Trainor LJ. 2006 Innateness, learning, and the difficulty of determining whether music is an evolutionary adaptation. *Music Percept.* **24**, 105–110. (doi:10.1525/mp.2006.24.1.105)

86. Snowdon CT. 2017 Vocal Communication in Family-Living and Pair-Bonded Primates. In *Primate Hearing and Communication* (eds RM Quam, MA Ramsier, RR Fay, AN Popper), pp. 141–174. Cham: Springer International Publishing. (doi:10.1007/978-3-319-59478-1_6)

87. Schruth DM. 2020 Musical calling as a behavior ancestral to all modern primates. *PsyArXiv*. (doi:10.31234/osf.io/mkze8)

88. Levréro F, Touitou S, Frédet J, Nairaud B, Guéry J-P, Lemasson A. 2019 Social bonding drives vocal exchanges in Bonobos. *Sci. Rep.* **9**, 711. (doi:10.1038/s41598-018-36024-9)

89. Fitch WT. 2005 The evolution of music in comparative perspective. *Ann. N. Y. Acad. Sci.* **1060**, 29–49. (doi:10.1196/annals.1360.004)

90. Marler PR, Krebs JR, Horn G. 1990 Song learning: the interface between behaviour and neuroethology. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **329**, 109–114. (doi:10.1098/rstb.1990.0155)

91. Balter M. 2010 Evolution of language. Animal communication helps reveal roots of language. *Science* **328**, 969–971. (doi:10.1126/science.328.5981.969)

92. Jarvis EDD. 2004 Learned birdsong and the neurobiology of human language. *Ann. N. Y. Acad. Sci.* **1016**, 749.

93. Lameira AR. 2017 Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neurosci. Biobehav. Rev.* **83**, 429–439. (doi:10.1016/j.neubiorev.2017.09.021)

94. Perlman M. 2017 Debunking two myths against vocal origins of language: Language is iconic and multimodal to the core. *Interact. Stud.* **18**, 376–401. (doi:10.1075/is.18.3.05per)

95. Marler P. 2000 Origins of music and speech: Insights from animals. In *The Origins of Music* (eds NL Wallin, B Merker, S Brown), pp. 31–48. Cambridge, MA: MIT Press.

96. Marler P, Dufty A, Pickert R. 1986 Vocal communication in the domestic chicken: I. Does a sender communicate information about the quality of a food referent to a receiver? *Anim. Behav.* **34**, 188–193. (doi:10.1016/0003-3472(86)90022-9)

97. Evans CS, Marler P. 1994 Food calling and audience effects in male chickens, Gallus gallus: their relationships to food availability, courtship and social facilitation. *Anim. Behav.* **47**, 1159–1170. (doi:10.1006/anbe.1994.1154)

98. Marler P, Dufty A, Pickert R. 1986 Vocal communication in the domestic chicken: II. Is a sender sensitive to the presence and nature of a receiver? *Anim. Behav.* **34**, 194–198. (doi:10.1016/0003-3472(86)90023-0)

99. Gyger M, Marler P. 1988 Food calling in the domestic fowl, Gallus gallus: the role of external referents and deception. *Anim. Behav.* **36**, 358–365. (doi:10.1016/S0003-3472(88)80006-X)

100. Karakashian SJ, Gyger M, Marler P. 1988 Audience effects on alarm calling in chickens (Gallus gallus). *J. Comp. Psychol.* **102**, 129–135. (doi:10.1037/0735-7036.102.2.129)

101. Gyger M, Marler P, Pickert R. 1987 Semantics of an Avian Alarm Call System: the Male Domestic Fowl, Gallus Domesticus. *Behaviour* **102**, 15–40. (doi:10.1163/156853986X00027)

102. Evans CS, Evans L, Marler P. 1993 On the meaning of alarm calls: functional reference in an avian vocal system. *Anim. Behav.* **46**, 23–38. (doi:10.1006/anbe.1993.1158)

103. Seyfarth RM, Cheney DL, Marler P. 1980 Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210**, 801–803. (doi:10.1126/science.7433999)

104. Macedonia JM. 2010 What is communicated in the antipredator calls of lemurs: Evidence from playback experiments with ringtailed and ruffed lemurs. *Ethology* **86**, 177–190. (doi:10.1111/j.1439-0310.1990.tb00428.x)

105. Hauser MD, Marler P. 1993 Food-associated calls in rhesus macaques (Macaca mulatta): I. Socioecological factors. *Behav. Ecol.* **4**, 194–205. (doi:10.1093/beheco/4.3.194)

106. Dittus WPJ. 1984 Toque macaque food calls: Semantic communication concerning food distribution in the environment. *Anim. Behav.* **32**, 470–477. (doi:10.1016/S0003-3472(84)80283-3)

107. Hauser MD, Teixidor P, Fields L, Flaherty R. 1993 Food-elicited calls in chimpanzees: effects of food quantity and divisibility. *Anim. Behav.* **45**, 817–819. (doi:10.1006/anbe.1993.1096)

108. Arnold K, Zuberbühler K. 2008 Meaningful call combinations in a non-human primate. *Curr. Biol.* **18**, R202-203. (doi:10.1016/j.cub.2008.01.040)

109. Arnold K, Zuberbühler K. 2006 Language evolution: semantic combinations in primate calls. *Nature* **441**, 303. (doi:10.1038/441303a)

110. Marler P, Pickert R. 1984 Species-universal microstructure in the learned song of the swamp sparrow (Melospiza georgiana). *Anim. Behav.* **32**, 673–689. (doi:10.1016/S0003-3472(84)80143-8)

111. Kroodsma DE. 1980 Winter wren singing behavior: a pinnacle of song complexity. *The Condor* **82**, 357–365. (doi:10.2307/1367556)

112. Kroodsma DE, Momose H. 1991 Songs of the Japanese population of the winter wren (Troglodytes troglodytes). *The Condor* **93**, 424–432. (doi:10.2307/1368959)

113. Van Horne B. 1995 Assessing vocal variety in the winter wren, a bird with a complex repertoire. *The Condor* **97**, 39–49. (doi:10.2307/1368981)

114. Ravignani A, Bella SD, Falk S, Kello CT, Noriega F, Kotz SA. 2019 Rhythm in speech and animal vocalizations: a cross-species perspective. *Ann. N. Y. Acad. Sci.* **1453**, 79–98. (doi:https://doi.org/10.1111/nyas.14166)

115. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830. (doi:10.1016/j.cub.2009.03.038)

116. Patel AD, Iversen JR, Bregman MR, Schulz I. 2009 Studying synchronization to a musical beat in nonhuman animals. *Ann. N. Y. Acad. Sci.* **1169**, 459–469. (doi:10.1111/j.1749-6632.2009.04581.x)

117. Schachner A, Brady TF, Pepperberg IM, Hauser MD. 2009 Spontaneous motor entrainment to music in multiple vocal mimicking species. *Curr. Biol.* **19**, 831–836. (doi:10.1016/j.cub.2009.03.061)

118. Hasegawa A, Okanoya K, Hasegawa T, Seki Y. 2011 Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Sci. Rep.* **1**, 120. (doi:10.1038/srep00120)

119. Fitch WT. 2013 Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Front. Syst. Neurosci.* **7**, 68. (doi:10.3389/fnsys.2013.00068)

120. Hattori Y, Tomonaga M, Matsuzawa T. 2013 Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Sci. Rep.* **3**, 1566. (doi:10.1038/srep01566)

121. Hattori Y, Tomonaga M, Matsuzawa T. 2015 Distractor Effect of Auditory Rhythms on Self-Paced Tapping in Chimpanzees and Humans. *PLOS ONE* **10**, e0130682. (doi:10.1371/journal.pone.0130682)

122. Hattori Y, Tomonaga M. 2020 Rhythmic swaying induced by sound in chimpanzees (Pan troglodytes). *Proc. Natl. Acad. Sci.* **117**, 936–942. (doi:10.1073/pnas.1910318116)

123. Payne K. 2000 The progressively changing songs of humpback whales: A window on the creative process in a wild animal. In (eds NL Wallin, B Merker, S Brown), pp. 135–150. Cambridge, MA: MIT Press.

124. Payne K, Tyack P, Payne R. 1983 Progressive changes in the songs of humpback whales (Megaptera novaeangliae): A detailed analysis of two seasons in hawaii. In *Communication and Behavior of Whales* (ed R Payne), pp. 9–57. Boulder, Colo.: Westview Press.

125. Green SR, Mercado E, Pack AA, Herman LM. 2011 Recurring patterns in the songs of humpback whales (Megaptera novaeangliae). *Behav. Processes* **86**, 284–294.

126. Eriksen N, Miller LA, Tougaard J, Helweg DA. 2005 Cultural change in the songs of humpback whales (Megaptera novaeangliae) from Tonga. *Behaviour* **142**, 305–328.

127. Noad MJ, Cato DH, Bryden MM, Jenner MN, Jenner KC. 2000 Cultural revolution in whale songs. *Nature* **408**, 537. (doi:10.1038/35046199)

128. Luther D, Baptista L. 2010 Urban noise and the cultural evolution of bird songs. *Proc. R. Soc. B Biol. Sci.* **277**, 469–473. (doi:10.1098/rspb.2009.1571)

129. Podos J, Huber SK, Taft B. 2004 Bird song: the interface of evolution and mechanism. *Annu. Rev. Ecol. Evol. Syst.* **35**, 55–87.

130. Miller G. 2011 *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature*. New York: Anchor Books.

131. Charlton BD. 2014 Menstrual cycle phase alters women's sexual preferences for composers of more complex music. *Proc. R. Soc. B Biol. Sci.* **281**, 20140403. (doi:10.1098/rspb.2014.0403)

132. Charlton BD, Filippi P, Fitch WT. 2012 Do women prefer more complex music around ovulation? *PLOS ONE* **7**, e35626. (doi:10.1371/journal.pone.0035626)

133. Marin MM, Schober R, Gingras B, Leder H. 2017 Misattribution of musical arousal increases sexual attraction towards opposite-sex faces in females. *PLOS ONE* **12**, e0183531. (doi:10.1371/journal.pone.0183531)

134. Madison G, Holmquist J, Vestin M. 2018 Musical improvisation skill in a prospective partner is associated with mate value and preferences, consistent with sexual selection and parental investment theory: Implications for the origin of music. *Evol. Hum. Behav.* **39**, 120–129. (doi:10.1016/j.evolhumbehav.2017.10.005)

135. Järvelä I. 2018 Genomics studies on musical aptitude, music perception, and practice. *Ann. N. Y. Acad. Sci.* **1423**, 82–91. (doi:10.1111/nyas.13620)

136. Trainor LJ. 2015 The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philos. Trans. R. Soc. B Biol. Sci.* **370**. (doi:10.1098/rstb.2014.0089)

137. Dunbar RIM. 2003 The origin and subsequent evolution of language. In *Language Evolution. Studies in the evolution of language.* (eds MH Christiansen, S Kirby), pp. 219–234. New York, NY: Oxford university Press. (doi:10.1093/acprof:oso/9780199244843.003.0012)

138. Dunbar RIM. 1996 *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

139. Dunbar RIM. 1993 Coevolution of neocortical size, group size and language in humans. *Behav. Brain Sci.* **16**, 681–735. (doi:10.1017/S0140525X00032325)

140. McComb K, Semple S. 2005 Coevolution of vocal communication and sociality in primates. *Biol. Lett.* **1**, 381–385. (doi:10.1098/rsbl.2005.0366)

141. Savage PE, Loui P, Tarr B, Schachner A, Glowacki L, Mithen S, Fitch WT. In press. Music as a coevolved system for social bonding. *Behav. Brain Sci.* , 1–36. (doi:10.1017/S0140525X20000333)

142. Kirby S. 2011 Darwin's musical protolanguage: an increasingly compelling picture. In *Language and Music as Cognitive Systems* (eds P Rebuschat, M Rohmeier, JA Hawkins, I Cross), pp. 96–102. Oxford: Oxford University Press. (doi:10.1093/acprof:oso/9780199553426.001.0010)

143. Fitch WT. 2011 The biology and evolution of rhythm: unravelling a paradox. In *Language and music as cognitive systems* (eds P Rebuschat, M Rohmeier, JA Hawkins, I Cross), pp. 73–95. Oxford: Oxford University Press. (doi:10.1093/acprof:oso/9780199553426.001.0009)

144. Brown S. 2000 The "musilanguage" model of music evolution. In *The Origins of Music* (eds NL Wallin, B Merker, S Brown), pp. 271–300. Cambridge, MA: MIT Press.

145. Baroni M. 2008 Music, musicality, 'musilanguage'. *Music. Sci.* **12**, 197–218. (doi:10.1177/1029864908012001091)

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

147. Fernald A, Kuhl PK. 1987 Acoustic determinants of infant preference for motherese speech. *Infant Behav. Dev.* **10**, 279–293. (doi:10.1016/0163-6383(87)90017-8)

148. Golinkoff RM, Can DD, Soderstrom M, Hirsh-Pasek K. 2015 (Baby)Talk to Me: The Social Context of Infant-Directed Speech and Its Effects on Early Language Acquisition. *Curr. Dir. Psychol. Sci.* **24**, 339–344. (doi:10.1177/0963721415595345)

149. Werker JF, McLeod PJ. 1989 Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Can. J. Psychol.* **43**, 230–246. (doi:10.1037/h0084224)

150. Mehr SA, Song LA, Spelke ES. 2016 For 5-Month-Old Infants, Melodies Are Social. *Psychol. Sci.* **27**, 486–501. (doi:10.1177/0956797615626691)

151. Mehr SA, Spelke ES. 2018 Shared musical knowledge in 11-month-old infants. *Dev. Sci.* **21**, e12542. (doi:10.1111/desc.12542)

152. Cirelli LK, Jurewicz ZB, Trehub SE. 2019 Effects of Maternal Singing Style on Mother–Infant Arousal and Behavior. *J. Cogn. Neurosci.* , 1–8. (doi:10.1162/jocn_a_01402)

153. Mehr SA, Krasnow MM, Bryant GA, Hagen EH. In press. Origins of music in credible signaling. *Behav. Brain Sci.* , 1–41. (doi:10.1017/S0140525X20000345)

154. Roeske TC, Tchernichovski O, Poeppel D, Jacoby N. 2020 Categorical Rhythms Are Shared between Songbirds and Humans. *Curr. Biol.* **30**, 3544-3555.e6. (doi:10.1016/j.cub.2020.06.072)

155. Bregman MR, Patel AD, Gentner TQ. 2012 Stimulus-dependent flexibility in non-human auditory pitch processing. *Cognition* **122**, 51–60. (doi:10.1016/j.cognition.2011.08.008)

156. Tierney AT, Russo FA, Patel AD. 2011 The motor origins of human and avian song structure. *Proc. Natl. Acad. Sci.* **108**, 15510–15515. (doi:10.1073/pnas.1103882108)

157. Bregman MR, Patel AD, Gentner TQ. 2016 Songbirds use spectral shape, not pitch, for sound pattern recognition. *Proc. Natl. Acad. Sci.* **113**, 1666–1671. (doi:10.1073/pnas.1515380113)

158. Ravignani A. 2018 Darwin, Sexual Selection, and the Origins of Music. *Trends Ecol. Evol.* **33**, 716–719. (doi:10.1016/j.tree.2018.07.006)

159. Mosing MA, Verweij KJH, Madison G, Pedersen NL, Zietsch BP, Ullén F. 2015 Did sexual selection shape human music? Testing predictions from the sexual selection hypothesis of music evolution using a large genetically informative sample of over 10,000 twins. *Evol. Hum. Behav.* **36**, 359–366. (doi:10.1016/j.evolhumbehav.2015.02.004)

160. Askin N, Mauskapf M, Koppman S, Uzzi B. 2020 Do women produce more novel work than men? Gender differences in musical creativity. *Work. Pap.*

161. Dunbar RIM. 2017 Group size, vocal grooming and the origins of language. *Psychon. Bull. Rev.* **24**, 209–212. (doi:10.3758/s13423-016-1122-6)

162. Grueter CC, Bissonnette A, Isler K, van Schaik CP. 2013 Grooming and group cohesion in primates: implications for the evolution of language. *Evol. Hum. Behav.* **34**, 61–68. (doi:10.1016/j.evolhumbehav.2012.09.004)

163. In press. Human grooming in comparative perspective: People in six small-scale societies groom less but socialize just as much as expected for a typical primate - Jaeggi - 2017 - American Journal of Physical Anthropology - Wiley Online Library. See https://onlinelibrary.wiley.com/doi/abs/10.1002/ajpa.23164 (accessed on 2 June 2021).

164. Dunbar RIM, Lehmann J. 2013 Grooming and social cohesion in primates: a comment on Grueter et al. *Evol. Hum. Behav.* **34**, 453–455. (doi:10.1016/j.evolhumbehav.2013.08.003)

165. Wyatt TD. 2015 The search for human pheromones: the lost decades and the necessity of returning to first principles. *Proc. R. Soc. B Biol. Sci.* **282**, 20142994–20142994. (doi:10.1098/rspb.2014.2994)

166. Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M. 1998 Increased auditory cortical representation in musicians. *Nature* **392**, 811–814. (doi:10.1038/33918)

167. Leipold S, Brauchli C, Greber M, Jäncke L. 2019 Absolute and relative pitch processing in the human brain: neural and behavioral evidence. *Brain Struct. Funct.* **224**, 1723–1738. (doi:10.1007/s00429-019-01872-2)

168. Wong PCM, Ciocca V, Chan AHD, Ha LYY, Tan L-H, Peretz I. 2012 Effects of Culture on Musical Pitch Perception. *PLOS ONE* **7**, e33424. (doi:10.1371/journal.pone.0033424)

169. Weisman RG, Ratcliffe L. 2004 Relative Pitch and the Song of Black-Capped Chickadees: Chickadees, like people, have a strong sense of relative pitch. These birds use skillful, precise pitch changes to advertise their quality and attract mates. *Am. Sci.* **92**, 532–539. (doi:10.1511/2004.50.948)

170. Hulse SH, Cynx J, Humpal J. 1984 Absolute and relative pitch discrimination in serial pitch perception by birds. *J. Exp. Psychol. Gen.* **113**, 38–54. (doi:10.1037/0096-3445.113.1.38)

171. Yin P, Fritz JB, Shamma SA. 2010 Do ferrets perceive relative pitch? *J. Acoust. Soc. Am.* **127**, 1673–1680. (doi:10.1121/1.3290988)

172. Stefanics G, Háden GP, Sziller I, Balázs L, Beke A, Winkler I. 2009 Newborn infants process pitch intervals. *Clin. Neurophysiol.* **120**, 304–308. (doi:10.1016/j.clinph.2008.11.020)

173. Trainor LJ, Austin CM, Desjardins RN. 2000 Is Infant-Directed Speech Prosody a Result of the Vocal Expression of Emotion? *Psychol. Sci.* **11**, 188–195. (doi:10.1111/1467-9280.00240)

174. Fernald A. 1993 Approval and disapproval: infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Dev.* **64**, 657–674.

175. Papoušek M, Bornstein MH, Nuzzo C, Papoušek H, Symmes D. 1990 Infant responses to prototypical melodic contours in parental speech. *Infant Behav. Dev.* **13**, 539–545. (doi:10.1016/0163-6383(90)90022-Z)

176. Malloch SN. 2000 Mothers and infants and communicative musicality. *Music. Sci.* **2**, 29–57. (doi:10.1177/10298649000030S104)

177. Winkler I, Háden GP, Ladinig O, Sziller I, Honing H. 2009 Newborn infants detect the beat in music. *Proc. Natl. Acad. Sci.* **106**, 2468–2471. (doi:10.1073/pnas.0809035106)

178. Clayton M, Jakubowski K, Eerola T, Keller PE, Camurri A, Volpe G, Alborno P. 2020 Interpersonal Entrainment in Music Performance: Theory, Method, and Model. *Music Percept.* **38**, 136–194. (doi:10.1525/mp.2020.38.2.136)

179. Pearce E, Launay J, MacCarron P, Dunbar RIM. 2017 Tuning in to others: Exploring relational and collective bonding in singing and non-singing groups over time. *Psychol. Music* **45**, 496–512. (doi:10.1177/0305735616667543)

180. Pearce E, Launay J, van Duijn M, Rotkirch A, David-Barrett T, Dunbar RIM. 2016 Singing together or apart: The effect of competitive and cooperative singing on social bonding within and between sub-groups of a university Fraternity. *Psychol. Music* **44**, 1255–1273. (doi:10.1177/0305735616636208)

181. Pearce E, Launay J, Dunbar RIM. 2015 The ice-breaker effect: singing mediates fast social bonding. *R. Soc. Open Sci.* **2**, 150221. (doi:10.1098/rsos.150221)

182. Weinstein D, Launay J, Pearce E, Dunbar RIM, Stewart L. 2016 Singing and social bonding: changes in connectivity and pain threshold as a function of group size. *Evol. Hum. Behav.* **37**, 152–158. (doi:10.1016/j.evolhumbehav.2015.10.002)

183. Hagen EH, Hammerstein P. 2009 Did Neanderthals and other early humans sing? Seeking the biological roots of music in the territorial advertisements of primates, lions, hyenas, and wolves. *Music. Sci.* **13**, 291–320. (doi:10.1177/1029864909013002131)

184. Hagen EH, Bryant GA. 2003 Music and dance as a coalition signaling system. *Hum. Nat.* **14**, 21–51. (doi:10.1007/s12110-003-1015-z)

185. Fritz T *et al.* 2013 Musical agency reduces perceived exertion during strenuous physical performance. *Proc. Natl. Acad. Sci.* **110**, 17784–17789. (doi:10.1073/pnas.1217252110)

186. Weisman O, Delaherche E, Rondeau M, Chetouani M, Cohen D, Feldman R. 2013 Oxytocin shapes parental motion during father-infant interaction. *Biol. Lett.* **9**, 20130828. (doi:10.1098/rsbl.2013.0828)

187. Feldman R, Weller A, Zagoory-Sharon O, Levine A. 2007 Evidence for a neuroendocrinological foundation of human affiliation: plasma oxytocin levels across pregnancy and the postpartum period predict mother-infant bonding. *Psychol. Sci.* **18**, 965–970. (doi:10.1111/j.1467-9280.2007.02010.x)

188. Gordon I, Zagoory-Sharon O, Leckman JF, Feldman R. 2010 Oxytocin and the development of parenting in humans. *Biol. Psychiatry* **68**, 377–382. (doi:10.1016/j.biopsych.2010.02.005)

189. Broad KD, Curley JP, Keverne EB. 2006 Mother-infant bonding and the evolution of mammalian social relationships. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 2199–2214. (doi:10.1098/rstb.2006.1940)

190. Kuhl PK. 2004 Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* **5**, 831–843. (doi:10.1038/nrn1533)

191. Ghazanfar AA, Liao DA, Takahashi DY. 2019 Volition and learning in primate vocal behaviour. *Anim. Behav.* **151**, 239–247. (doi:10.1016/j.anbehav.2019.01.021)

192. Takahashi DY, Liao DA, Ghazanfar AA. 2017 Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. *Curr. Biol.* **27**, 1844-1852.e6. (doi:10.1016/j.cub.2017.05.004)

193. Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, Ghazanfar AA. 2015 The developmental dynamics of marmoset monkey vocal production. *Science* **349**, 734–738. (doi:10.1126/science.aab1058)

194. Toro JM, Crespo-Bojorque P. 2017 Consonance Processing in the Absence of Relevant Experience: Evidence from Nonhuman Animals. *Comp. Cogn. Behav. Rev.* **12**. (doi:10.3819/CCBR.2017.120004)

195. Freeberg TM. 2006 Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychol. Sci.* **17**, 557–561. (doi:10.1111/j.1467-9280.2006.01743.x)

196. Blumstein DT, Armitage KB. 1997 Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *Am. Nat.* **150**, 179–200. (doi:10.1086/286062)

197. Furrer RD, Manser MB. 2009 The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *Am. Nat.* **173**, 400–410. (doi:10.1086/596541)

198. Wilkinson GS. 2003 Social and vocal complexity in bats. In *Animal social complexity: Intelligence, culture, and individualized societies* (eds FBM de Waal, PL Tyack), pp. 322–341. Cambridge, MA: Harvard University Press.

199. Freeberg TM, Dunbar RIM, Ord TJ. 2012 Social complexity as a proximate and ultimate factor in communicative complexity. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1785–1801. (doi:10.1098/rstb.2011.0213)

200. Peckre L, Kappeler PM, Fichtel C. 2019 Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behav. Ecol. Sociobiol.* **73**, 11. (doi:10.1007/s00265-018-2605-4)

201. Rebout N *et al.* 2020 Tolerant and intolerant macaques show different levels of structural complexity in their vocal communication. *Proc. R. Soc. B Biol. Sci.* **287**, 20200439. (doi:10.1098/rspb.2020.0439)

202. Pollard KA, Blumstein DT. 2012 Evolving communicative complexity: insights from rodents and beyond. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1869–1878. (doi:10.1098/rstb.2011.0221)

203. Elowson AM, Snowdon C, Lazaro-Perea C. 1998 Infant 'Babbling' in a nonhuman primate: Complex vocal sequences with repeated call types. *Behaviour* **135**, 643–664. (doi:10.1163/156853998792897905)

204. Snowdon C, Elowson AM. 2001 'Babbling' in pygmy marmosets: Development after infancy. *Behaviour* **138**, 1235–1248. (doi:10.1163/15685390152822193)

205. Burkart J, Martins EG, Miss F, Zürcher Y. 2018 From sharing food to sharing information: Cooperative breeding and language evolution. *Interact. Stud.* **19**, 136–150. (doi:10.1075/is.17026.bur)

206. Hrdy SB, Burkart JM. 2020 The emergence of emotionally modern humans: implications for language and learning. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190499. (doi:10.1098/rstb.2019.0499)

207. Gustison ML, le Roux A, Bergman TJ. 2012 Derived vocalizations of geladas (Theropithecus gelada) and the evolution of vocal complexity in primates. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1847–1859. (doi:10.1098/rstb.2011.0218)

208. Richman B. 1978 The synchronization of voices by gelada monkeys. *Primates* **19**, 569–581. (doi:10.1007/BF02373317)

209. Richman B. 1987 Rhythm and melody in gelada vocal exchanges. *Primates* **28**, 199–223. (doi:10.1007/BF02382570)

210. Suire A, Raymond M, Barkat-Defradas M. 2019 Male Vocal Quality and Its Relation to Females' Preferences. *Evol. Psychol.* **17**, 1474704919874675. (doi:10.1177/1474704919874675)

211. Tognetti A, Durand V, Barkat-Defradas M, Hopfensitz A. 2020 Does he sound cooperative? Acoustic correlates of cooperativeness. *Br. J. Psychol.* **111**, 823–839. (doi:10.1111/bjop.12437)

212. Castro M, L'héritier F, Plailly J, Saive A-L, Corneyllie A, Tillmann B, Perrin F. 2020 Personal familiarity of music and its cerebral effect on subsequent speech processing. *Sci. Rep.* **10**, 14854. (doi:10.1038/s41598-020-71855-5)

213. Stewart-Williams S, Thomas AG. 2013 The Ape That Thought It Was a Peacock: Does Evolutionary Psychology Exaggerate Human Sex Differences? *Psychol. Inq.* **24**, 137–168. (doi:10.1080/1047840X.2013.804899)

214. Feinberg DR. 2008 Are human faces and voices ornaments signaling common underlying cues to mate value? *Evol. Anthropol.* **17**, 112–118. (doi:10.1002/evan.20166)