

The role of musical pitch in long-distance defensive signaling

Piotr Podlipniak 

Department of Musicology, Adam Mickiewicz University, ul. Uniwersytetu Poznańskiego7, Poznań, Poland

Correspondence

Piotr Podlipniak, Department of Musicology, Adam Mickiewicz University, ul. Uniwersytetu Poznańskiego7, Poznań, 61-614, Poland.
Email: podlip@amu.edu.pl

Funding information

National Science Center, Poland

Abstract

Many species use conspicuous sound signaling as a defense strategy against conspecifics and heterospecifics. This has prompted some scholars to hypothesize that musical displays can serve as a defensive signaling mechanism. It has been suggested that to achieve this function, hominins may have used synchronized acoustic signals, which then became the roots of musical rhythm. Another important component of music—pitch—has been proposed to result either from other selective pressures or as a byproduct of certain biological constraints. This paper aims to add to the idea of the defensive function of music, by looking at the role of pitch matching in the long-distance deterrence of conspecifics between singing hominins. Since the production and recognition of musical pitch depend on many abilities, their evolution in the context of this possible defensive function is discussed. In particular, the emergence and role of culturally variable musical pitch elements as hallmarks of group identity are examined when considering selective pressures that occurred from increasing hominin social complexity. In contrast to current monofunctional explanations, this proposal highlights a possible reinforcement between other adaptive functions of musical pitch related to the hominin social niche, which ultimately created a synergistic loop and accelerated the evolution of musicality.

KEYWORDS

biomusicology, defense signaling, intraspecific defense, long-distance calls, musicality, musical pitch

INTRODUCTION

Pitch is a perceptual sensation that usually accompanies our auditory experience of harmonic sounds. These are complex sounds whose partials (harmonics in this case) are integer multiples of a fundamental frequency (F_0), and they can be produced by both animate and inanimate sources. While humans likely share the ability to recognize harmonic sounds with other primates,¹ some differences in cortical specialization for pitch processing have been observed between *Homo sapiens* and other primates,² which may be due to the significant role that pitch plays in human-specific forms of auditory

communication. Among them, music is one domain in which the sensation of pitch appears to be especially crucial. Although pitch is also present in almost all forms of human vocalizations such as crying, laughter, and speech, the experience of pitch in music consists of certain additional and specific cognitive properties such as discreteness^{3,4} and circularity.⁵ While pitch discreteness may initially appear to be an objective characteristic of stimuli rather than a cognitive construct, there is ample evidence suggesting otherwise. For example, in the speech-to-song transformation,⁶ where looped speech starts to be perceived as a song, the same continuous changes in a vowel's F_0 in speech are interpreted as discrete pitches only because of their repetition and

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Annals of the New York Academy of Sciences* published by Wiley Periodicals LLC on behalf of The New York Academy of Sciences.

not because of the stability of F_0 , which usually differentiates songs and music from speech.^{7,8} This same effect is observed in the sounds-to-music transformation.⁹ Such discrete sensations are then incorporated into the circular dimension of pitch, which is characterized in Western music (and many other musical cultures) by the repetition of pitch classes in each octave. While particular discrete pitches are conventionalized, the tendency to discretize pitch in music seems to be a general cross-cultural phenomenon¹⁰ which suggests the existence of cognitive constraints related to this feature. Additionally, the experience of pitch in tonal music, which dominates musical cultures globally,¹¹ is accompanied by subtle sensations of stability,¹² resulting in the interpretation of pitch sequences in terms of pitch hierarchies.

The existence of such specific pitch sensations in music leads to the question of their origin. As music appears to be a rather wasteful expenditure, some researchers have proposed that musical pitch is a byproduct of other adaptive abilities such as the language faculty,¹³ auditory scene analysis,¹⁴ or other nonmusical biological constraints resulting from the features of speech vocalizations.¹⁵ However, the fact that musical pitch is universal^{11,16,17} and is part of music-specific mental representations strongly counters any byproduct explanations for its origins. Therefore, many hypotheses indicate possible adaptive values of musicality^{18–34} as a set of abilities that allow the recognition and production of music.^{35,36} Yet, even though the ability to experience musical pitch belongs to one of the crucial components of musicality, most of these hypotheses focus on the adaptive value of musicality in general, and only some look for the specific functions related to particular musical traits.^{32,37–39} However, while the origin of musical rhythm that is based on auditory–motor synchronization has often been explained in terms of social bonding⁴⁰ or signaling coalition quality,²⁰ musical pitch seems to have received less attention in this respect.

Admittedly, this topic has recently begun to attract some interest among scholars. It has been suggested, for instance, that chorusing could have become an important part of early hominins' vocalizations,⁴¹ serving the social bonding function as a kind of vocal grooming.⁴² Although the evolution of musical pitch might have been related to social consolidation, the vocal grooming hypothesis does not explain why musical pitch has specific features such as discreteness, hierarchy, and so on. Similarly, the idea that musical pitch could have been a fitness indicator as part of a sexual display²⁶ leaves the question of the origins of the specific characteristics of musical pitch unanswered. It has been also proposed that the evolution of the melodic part of music originated as a kind of attention signaling in an evolutionary arms race resulting from differences in fitness interests between parents and offspring.^{43,44} Considering that the pitch structure of infant-directed singing and music⁴⁵ differs from the use of pitch in adult-directed music,^{46–48} this explanation seems at least partially incomplete. In this paper, an alternative view is proposed, in which the pitch part of musicality was shaped by additional selective pressures, including the long-distance deterrent of conspecific outsiders with malicious intent (see Figure 1). Unlike previous proposals that either neglect the defensive function of music^{18,19,24–26,29} or point to the role of rhythm in fulfilling this function,^{20,22,44,34} the hypothesis

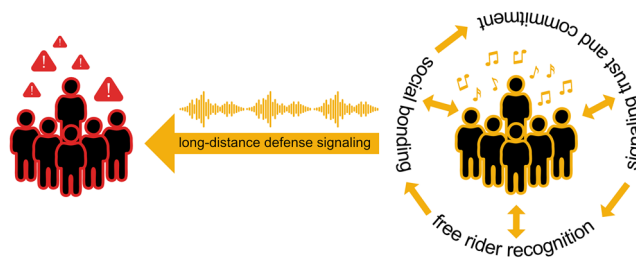


FIGURE 1 The functions of musical pitch. The same pitch sequence can act as a defense signal or as a tool for social bonding, free rider recognition, and signaling trust and commitment depending on the in-group or out-group perspective. Red—out-group hominins; Yellow—in-group hominins.

proposed here emphasizes the importance of pitch as a part of defense signaling. According to this novel view, the adaptive values of musicality do not have to be mutually exclusive,^{49,50} not only because they may be related to different elements of musicality, but also because they may refer to different stages of the evolution of given components of musicality. Therefore, the possible adaptive function of pitch that is proposed here should be considered as an addition to the set of functions such as social bonding,²⁴ free rider recognition,³³ and signaling commitment that may have contributed to the selection of the pitch part of human musicality.

MUSIC AS A DEFENSE SIGNAL

The idea that music (and dance) could have originated as a hominin defense strategy against predators^{51,21,52–55} is based on the fact that many species use conspicuous vocalizations as alarm and aposematic signals. In fact, people from some tribal cultures still use musical behavior such as singing and drumming to deter wild animals.⁵³ Moreover, some elements of vocalizations can act as cues of aggression, body size, group size, and so forth,^{56,57} which may influence the behavior of predators. It has been observed, for instance, that lions are able to recognize the numerosity of conspecific intruders based on their roaring, which can result in the decision whether to approach them or not.⁵⁸ Although *H. sapiens* is currently the most feared super predator,⁵⁹ which may imply the lack of a need for any defense strategy, in the Early Pleistocene our ancestors were most likely active or confrontational scavengers,^{60–62} while not excluding being hunters too,⁶³ which involved competition with other predators or scavengers such as lions and giant hyenas for carcasses.³⁴ As the size of the hominin group could have given an advantage in this competition,⁶⁰ one of the key factors influencing a lion's or hyena's decision to fight or not for carcasses could have been the cue for the size of the hominin group. Since signals can evolve from cues,^{64,65} the presence of warning vocalizations in primates, which can additionally convey cues of aggression and group size, makes the hypothesis of aposematic signaling as the roots of music likely. Both Hagen and Jordania indicated the main feature of music is most likely to deter these predators is rhythm. This assumption is inferred from observations that many acoustic

aposematic signals are synchronized displays, and musical rhythm is based on the human ability to synchronize motor activity, including vocalizations, with auditory stimuli. It is worth mentioning in this context that according to Hagen,²¹ the deterrent function of music does not exclude, but complements the credible signaling hypothesis,^{44,20} which assumes that musical rhythm evolved as a signal of coalition quality (group size and its cooperation ability as well as coalition strength), whereas melody evolved as a signal of parental attention.⁴⁴ By contrast, for Knight and Lewis,⁵³ pitch variations could have been used as a defense strategy to deter approaching lions by giving them a misleading sense of hominin group size. Although both musical rhythm and musical pitch are human-specific signals, which may raise doubts about their recognition by other species, the acquisition of information from heterospecific signals is a common phenomenon in nature^{66–68} and can be both inherited⁶⁹ and learned,^{70,71} which makes the defensive hypotheses worth exploring.

MUSICAL PITCH AS AN INTRASPECIFIC DETERRENT

The origins of hominins are closely linked to the rise of terrestriality, which exposed early humans to greater predation risks and, consequently, intensified selective pressures that favored cooperative defense strategies.^{21,72,73} This collective defense, likely aimed at protecting against large carnivores, would have included the use of stones, sticks, and branches—especially those with sharp thorns⁷⁴—as weapons, alongside physical displays and loud vocalizations.²² This strategy would have been further supported by an upright posture linked to the evolution of bipedality, which may have helped scare away predators.⁵⁵ The threat of predation persisted through the emergence of the first *Homo* genus members, suggesting that collective defensive behaviors remained crucial to our ancestors' survival.⁷⁵ Given that both predation pressure and increasing social intricacy are thought to significantly influence communication complexity,^{76–78} it seems highly plausible that hominin vocalizations evolved to incorporate the specific use of harmonic sound as an aposematic signal. In this context, protomusical behaviors that functioned as interspecific aposematic signals may have acted as a catalyst for the emergence of protomusical, intraspecific defense signaling, characterized by more complex rhythmic and pitch structures.

Ethnomusicologists have explored for decades the relationship between society and music in terms of their structures and functions.^{79,80} Specifically, they have observed that the size and social composition of a musical group can correlate with specific musical characteristics, as exemplified by the organization of musical ensembles in Pygmy polyphony.⁸¹ Such a correlation may reflect the proposed causal link between human sociality and the evolution of musicality.⁸² Intergroup competition may have been one of the factors that influenced the tendency of hominins to express social traits through musical structure. After all, when well-consolidated groups of hominins became a threat, aposematism required more than just the display of numerosity and aggression—it demanded the signaling

of coalition strength.²⁰ Of course, this scenario does not necessarily exclude the possibility that, in primates, interspecific defense signaling evolved after intraspecific defense signaling. In fact, the need to deter conspecifics has been a significant factor in the evolution not only of hominins but also of other primates as intergroup aggression is a commonly observed behavior.⁸³ Although intergroup competition among hominins remains a topic of debate,^{84–87} with a potential inclination toward intergroup cooperation,⁷² the limited availability of resources could have increased the risk of intergroup conflicts, as is believed to have occurred in the competition between Neanderthals and *H. sapiens* groups.^{88,89}

As in the case of predation, the interest of conspecific intruders is to avoid fighting if it is costly and ineffective. Also, signalers producing signals is often less costly than fighting with individuals from a competing group.⁹⁰ Moreover, the ability to receive this information from a long distance allows potential intruders to avoid the risk of entering the territory of a foreign group. Therefore, a credible signal of group strength should also act as an effective defense tool against conspecifics, giving benefits for both senders and receivers of this signal. However, together with increasing social intelligence among hominins, the scope of potentially dangerous interactions between conspecifics would have increased.⁹¹ As a result, not only group strength, but also other factors such as the strength of bonds between group members, the size and composition of the group, the identity of a formerly met rival, and the group's experience with cooperation and combat would have become important knowledge to acquire that could have influenced the decision to attack or abandon an attack against a specific group.⁹⁰ In such circumstances, what was enough to deter heterospecific predators might not have been sufficient to avoid attack by conspecifics. Moreover, interactions with heterospecific and conspecific outsiders differ in that predators are always a threat, whereas conspecific outsiders can pose both a threat or a benefit resulting from cooperation⁹⁰ (e.g., by being part of a new coalition or as a betrayer of a rival group), which can extend the scope of possible behavioral reactions. All these new challenges could have led to the elaboration of defense strategies including the use of acoustic defense signals.

Although many primates use vocalizations as territorial defense signals which are directed toward conspecifics, these signals are usually instinctive calls^{92,93} sometimes subject to flexibility in use^{94,95} or even characterized by certain developmental plasticity.⁹⁶ In contrast, the most probable important advantage of hominin sound communication that could have been used as defense signaling directed against conspecific intruders was the fact that it was composed of more elaborate culture-specific parts. In the case of a perpetually changing social environment such as in the composition of groups and the new skills acquired by group members, the culture-specific elements of signals could have become an important source of information about ongoing changes. After all, such features like a newly invented sound pattern, an improved performance in terms of synchronization and/or pitch matching, and the addition of a new performer, carried an extra set of information to be inferred by conspecifics. It is of course a risk that in the case of the increasing complexity of signals some of the message can become obscured, which can cause a

tendency to increase redundancy of the signal.⁹⁷ In fact, music is characterized by immanent repetitions⁹⁸ that may be an effect caused by this risk. As our closest relatives, chimpanzees use both drumming⁹⁹ and vocalization^{100,101} in acoustic signaling, it is therefore reasonable to assume that similar forms of communication were accessible to hominins. However, while the use of percussion tools usually limits produced sounds to rhythmic patterns and makes the sound features dependent on the acoustic properties of percussion instruments,¹⁰² the use of vocalizations expands the scope of encoded information by adding other spectral features into a means of communication. Expanding the culture-specific elements of the signaling system on these new possible units based on spectral distinctions (employing timbral and pitch contrasts¹⁰³) increases the combinatoriality of this system. But did the increased complexity of the social system and the new defense challenges it brought contribute to the origins of music? Was the role of pitch in music somehow related to defense signaling?

A TANGLE OF ULTIMATE EXPLANATIONS FOR THE ORIGINS OF MUSICAL PITCH

Looking for one adaptive function of music seems to be delusive when considering that musicality is composed of many different abilities that did not evolve at the same time. Even the term *musical pitch* may actually refer to a set of biologically meaningful but different signals that may depend on different abilities. One possible way to identify the adaptive functions of the ability to recognize and produce musical pitch is to compare similar abilities and their adaptive functions among species other than humans.¹⁰⁴ Many abilities that allow for the perception and production of pitch are in fact present in several nonhuman animals.¹⁰⁵ For example, it has been identified among certain avian¹⁰⁶ and mammalian¹⁰⁷ species. Pitch matching has been observed in Japanese macaques (*Macaca fuscata*),¹⁰⁸ while harmonic intervals have been recognized in the songs of the hermit thrush (*Catharus guttatus*),¹⁰⁹ the great tit (*Parus major*),¹¹⁰ and the musician wren (*Cyphorhinus arada*).¹¹¹ Similarly, the ability to recognize octave equivalence—a sensation of pitch affinity as a result of the perceptive fusion of sounds with a F_0 ratio of 2:1^{112,113}—has been observed in dolphins (*Tursiops truncatus*)¹¹⁴ and rhesus macaques (*Macaca mulatta*).¹¹⁵

Yet, while we share certain features of musicality with other species, none produces music in its full human form.³⁶ Likewise, as far as we know, no nonhuman species possess the entire set of abilities mentioned above that characterize the part of musicality that enables the recognition and production of musical pitch. For instance, Wagner and Hoeschele have shown that even the four abilities they identified as crucial for the spectral aspects of human musicality—namely, vocal learning, harmonic clarity in vocalizations, differing vocal ranges, and simultaneous vocalization and duetting—have not been observed together in any other species.¹⁰⁴ Yet, in order for humans to use musical pitch, it is necessary to have all these abilities together. Although many abilities jointly compose the pitch part of musicality, the origin of each of them, as well as their primary adaptive functions, might not

have necessarily been related to musical behavior.¹¹⁶ In other words, some of these abilities may have evolved due to the adaptive value of protomusical behavior, while others might have evolved earlier due to different adaptive functions and were only later exapted and incorporated into musicality. For instance, vocal learning could have initially served an important function in broadening the scope of sounds used to communicate propositional meaning.¹¹⁷ A good example of such a strategy is the intentional imitation of low frequencies. This strategy could have been an adaptive behavior of communicating a broad meaning related to largeness due to the frequency code¹¹⁸—the remnants of which are probably still present in sound symbolism.¹¹⁹

Although the volitional vocal control of F_0 may have originally evolved as a tool for protoverbal communication through sound symbols,¹¹⁷ it might later have been exapted to intentionally control the pitch aspect of affective prosody. This process could have formed the basis of musical pitch as it took on a new role in deterring conspecifics. It is also possible that the volitional control of F_0 could have been used in order to communicate other meanings such as dominance, strength, masculinity, or femininity, depending on the social context.¹²⁰ A similarly intricate evolutionary path could have led to the emergence of the other abilities connected with musical pitch.²⁴ However, taking into account the variety of physiological mechanisms involved in the production and processing of musical pitch,^{121–125} the evolution of each of them as the result of one adaptive function seems to be unlikely. It is also possible that the convergent or parallel evolution of some abilities could have occurred in response to different selective pressures.¹²⁶ Therefore, octave equivalence, for instance, could have evolved independently in dolphins and primates as a tool to enhance communication by means of harmonic sounds, but in response to different functions of such communication.

Similarities between sounds composing octave could be used to restrict the number of units in a communicative system, or could have just been used for the expression of similarity in an iconic way. It is therefore presumable that octave equivalence among *H. sapiens* evolved due to additional selective pressures, for example, those related to social functions.¹²⁷ This does not exclude the possibility that the tendency to recognize pitches that share many overtones as similar may be a trait that evolved long before hominins.¹⁰⁴ Sexual dimorphism in voice pitch among humans, which amounts to approximately one octave, is significantly greater than that observed in other apes.¹²⁸ The hypothesis that sexual selection alone drove the evolution of octave equivalence appears to be insufficient,¹²⁷ particularly given that the F_0 difference between male and female voices typically diminishes with the transition to monogamy.¹²⁸ Since monogamy is the predominant mating system in *H. sapiens*, such a pronounced disparity in voice pitch between men and women likely stems from additional selective pressures. For instance, the act of male and female hominins singing together in octaves may have conferred an advantage in long-distance group signaling (see the section “The role of timbre, pitch, heterophony, and vibrato in signaling group size”). Thus, while the initial emergence of voice dimorphism in our ancestors was most likely driven by sexual selection,¹²⁹ the pitch difference between men and women eventually became a constraint influencing the evolution of vocal defense signal-

ing. To achieve pitch matching and overcome this constraint, singing in octaves may have proved to be the most effective solution.

The intraspecific defensive function of musical pitch mentioned above can be attributed to the use of discrete pitch patterns and the evolution of sensitivity to them.³⁰ From this perspective, pitch matching between singers can act as a credible signal of group size and cohesion. Such a signal decoded by the in-group members probably served a different function than that detected by out-group members, thus being attractive for the former and aversive for the latter. After all, the same signal can cause different reactions depending on its usefulness for a given individual. For example, the distress signal of a baby crying motivates caregivers to care,^{130,131} while it encourages crocodiles to prowl for easy prey.¹³²

A similar effect could have impacted the evolution of musical pitch among conspecifics. Different context-based meanings of the same signal (submission and subordination) have been observed in primates,¹³³ which makes it likely that a similar mechanism may have occurred in hominins using pitch matching as a form of collective vocalization. The defensive function of pitch patterns could have then been enhanced by a cultural variability of these signals as the emblems of particular groups, becoming at the same time a kind of friend or foe identification system. Such a strategy, often named password signaling, has been observed in many vocal learners such as cetaceans^{134,135} and birds.^{136–138} Fitch has hypothesized that password signaling could have been a reason for the evolution of human vocal learning.¹³⁹ According to Fitch, the human ability to recognize natives via a language accent speaks to this hypothesis. However, when taking into account that the recognition of a foreign accent is mainly influenced by speech rhythm,¹⁴⁰ which is based on amplitude modulation,¹⁴¹ password signaling by means of speech seems to be restricted to short distances (see the next section). In contrast, the changes of F_0 that are crucial for the recognition of pitch seems to be a better choice as a means to advertise group belonging over long distances. In noisy conditions, vowels—being harmonic sounds—are more easily recognized over long distances than consonants.¹⁴² The advantage of harmonic sounds in long-distance communication is further supported by the fact that many cultures use whistle languages, composed entirely of harmonic sounds, as tools for such communication.¹⁴³

Hagen and Hammerstein have suggested that musical variability could have evolved due to the need for communal identity signaling as part of territorial advertisement.³⁴ They have also emphasized the crucial role of vocal learning in this task. Therefore, the implicit learning of a culture-specific musical pitch system by contemporary humans may be a canalized remnant of using musical pitch as a defense signal. Above all, what is an easy task for us today could have been a greater challenge for our ancestors. For hominins with restricted vocal learning, adjusting to any pitch pattern must have been a strenuous task not only because of the demands on memory, but also because of the need to be in tune with a particular group-specific tuning. Being time-consuming and strenuous on the one hand, and adaptive as a signal of group consolidation that defends against hostile out-group conspecific attacks on the other, this sensitivity to tuning could have become an instinctive learning strategy. Instinctive learning, being a type of learning based on innate learning preferences which guide and facilitate

the learning process,¹⁴⁴ may have given an advantage to individuals by confirming belonging and commitment to the group as well as deterring out-group hostile conspecifics. After this evolution of sensitivity to tuning, the process of an arms race of learning abilities between individuals may have continued.³³ The new challenge, however, was pitch distribution.³² In both cases of sensitivity to tuning and sensitivity to pitch distribution, the experience of well-matched pitches could have caused positive emotional reactions in group members singing together. In contrast, such well-matched patterns could have triggered awe and fear in out-group conspecifics, serving as an effective defense strategy.

HARMONIC SOUNDS AS DEFENSE SIGNALS IN NONHUMAN TERRESTRIAL ANIMALS

As fighting is often more costly than sound production, animals often use sounds as defense signals aimed at both conspecifics and heterospecifics.^{90,145} Some of these behaviors can be very conservative, whereas others can evolve independently. As convergent evolution provides a special opportunity to recognize common ecological conditions that could have contributed to the evolution of similar traits among species not closely related,¹⁴⁶ it is useful to look at the defense signals used by different terrestrial animals. After all, similar defense signals used by unrelated species could have been caused by similar selective pressures. Many examples of defensive sound signals consist of broadband sounds. Energy is distributed over a large range of the spectrum in growling by the domestic cat (*Felis silvestris catus*),¹⁴⁷ dingo (*Canis lupus dingo*),¹⁴⁸ domestic dog (*Canis familiaris*),¹⁴⁹ Indian wolf (*Canis lupus pallipes*),¹⁵⁰ and so forth. Also, hissing by the domestic goose,¹⁵¹ blue tit (*Cyanistes caeruleus*),¹⁵² domestic cat (*F. silvestris catus*),¹⁴⁷ puff adder (*Bitis arietans*),¹⁵³ bumblebee (*Bombus terrestris*),¹⁵⁴ and Asian dwarf honeybee (*Apis florea*)¹⁵⁵ consists of such sounds. In fact, broadband sounds have many advantages in respect to defense functions. First, the wide distribution of energy on the partials of different frequencies provides enough redundancy that would be hard to ignore by the recipient. Second, broadband sounds are difficult to be masked entirely by environmental sounds. Finally, such a wide distribution of acoustic energy makes the signal accessible to various species that have different auditory frequency ranges. However, if a signal is to act over a long distance, its effectiveness depends on its robustness of sound propagation in the air. In this respect, broadband sounds are quite poor as they are susceptible to scattering and attenuation.¹⁵⁶

In contrast, harmonic sounds, especially at low frequencies, are much more effective at propagation in the air at longer distances.^{157–159} Therefore, terrestrial animals usually use harmonic sounds in sound communication whenever the potential recipients of the signals are located at a large distance (see Figure 2). This observation led to the development of the acoustic adaptation hypothesis, which posits that long-distance vocalizations evolve and persist because they are optimally transmitted within the specific habitats animals inhabit.¹⁶⁰ One prediction of this hypothesis is that in closed habitats, vocalizations tend to be longer, lower in frequency,

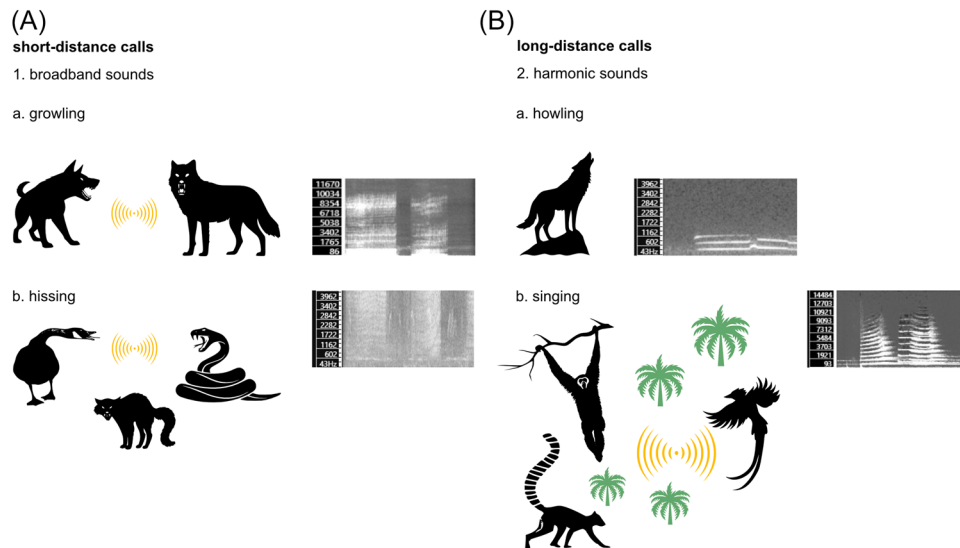


FIGURE 2 The two types of vocal signaling: (A) short-distance calls: a—growling (the spectrogram of a dog growl), b—hissing (the spectrogram of a hissing goose); (B) long-distance calls: a—howling (the spectrogram of a wolf's howl), b—singing (the spectrogram of an indri song).

and less repetitive, whereas in open habitats, the opposite pattern is expected. Although a computer simulation has supported the validity of this hypothesis,¹⁶¹ more recent meta-analyses of empirical studies on amphibians, birds, and mammals have not provided consistent support.¹⁶²

Nevertheless, the use of harmonic sounds in long-distance signaling is often observed among terrestrial animals. This rule seems to apply to defense signals as well. Of course, harmonicity is not the sole factor influencing the effectiveness of signal propagation, with sound intensity and background noise¹⁶³ ranking among the most significant. Therefore, the use of harmonic sounds in this type of signaling cannot be regarded as an absolute rule. A need to deter intruders from long distances is especially important in territorial species.^{164,165} After all, in order to sustain exclusive access to resources such as shelter, food, and mates, it is crucial to determine and mark the boundaries of the inhabited territory.^{166–168} According to Hagen and Hammerstein, the most common form of acoustic territorial advertisement is vocalization.³⁴ In fact, vocalizations composed of harmonic sounds are used as a territorial defense by many species of birds^{169–171} and mammals.^{150,172–177} Interestingly, harmonic sounds are also used in emergency sirens, which has led to the speculation about the possible role of the human predisposition to an attentional response to howling wolves and the invention of the siren.¹⁷⁸ But this response can be the result of an evolutionarily older sensitivity to loud harmonic sounds. Early species of the *Homo* genus were probably cooperative hunters,¹⁷⁹ which makes them similar to wolves, lions, and hyenas in that respect. Since all these animals advertise their territory by means of loud pitched vocalizations such as howling, roaring, and whooping, respectively, one would suspect that hominins also used defense signaling that could have resembled elements of these vocalizations.³⁴

Among mammals that use harmonic sounds in their long-distance calls, there are also primates such as marmosets,^{180–182} tamarins,^{182,183} titi monkeys,^{184,185} langurs,¹⁸⁶ indris,¹⁸⁷ tarsiers,¹⁸⁸

and gibbons.^{189,28} The ubiquity of long-distance vocalizations among primates suggests that human musicality has its roots in this behavior.²⁸ A common form of vocalization observed among these primates is duetting. In exploring the evolution of human musicality, researchers have emphasized the role of synchronization^{190,191} and pair bonding²⁸ in relation to primate duetting. It is worth mentioning, however, that minimal volitional control of pitch has been observed among many primates,^{192–194} including gibbons,¹⁹⁵ which opens the door to the possibility that some elements of vocal learning related to human singing could have been inherited from the common arboreal ancestor of humans and apes.^{55,54} Although the correlation between primate duetting and monogamy has been interpreted as a strong indication for the role of sexual selection in the evolution of primate singing,^{186,190} it should not be ignored that primate songs are also used as territorial defense signals,^{185,196,197} which speaks for the defense hypothesis of music origin. It should also not be ignored that harmonic sounds are an important element of these vocalizations, which prompts us to consider the role of pitch as an important factor in our hominin ancestor's defensive signaling over long distances, even though some primates, such as chimpanzees (*Pan troglodytes*), do not sing.¹⁹⁸

Lack of singing does not mean, however, that chimps do not use harmonic sounds in their vocalizations. Indeed, they are present in many types of chimpanzee vocalizations such as in laughter, screams, hoo calls, pant-hoots, barks, squeaks, as well as in some types of grunts and whimpers.¹⁰¹ Harmonic sounds are also an important component of affective prosody,¹⁹⁹ elements of which are present in the vocalizations of many mammalian species,^{200,201} including the aforementioned vocalizations of chimpanzees. For many scholars affective prosody seems to be a good candidate to be a direct precursor of both speech and singing in humans.^{31,202–206} Perhaps affective prosody also gave rise to the most complex vocalization specific to chimpanzees and bonobos (*Pan paniscus*), that is, pant-hooting—a form of long-

distance communication characterized by a sequence of hoo calls and screams.¹⁰¹ Pant-hooting can also be occasionally combined with rough grunts, which suggests volitional control over its production,²⁰⁷ and seems to develop later in ontogenesis in comparison to other types of chimp calls.¹⁰¹ Although territorial defense is not the only function proposed for pant-hooting,²⁰⁸ the fact that social functions such as signaling social bonds²⁰⁹ or social status,²¹⁰ attracting conspecifics to food,²¹¹ and facilitating fusion,²¹² have been indicated among its many different functions and suggests that chimpanzee vocalizations are functionally similar to those of other primates. This means that long-distance hominin calls could have also been under selective pressure related to conspecific social environments. An important difference between chimpanzee and *H. sapiens* long-distance vocalizations, which can be a result of this selective pressure, is the more elaborate cultural variability of the latter.

THE AFFILIATIVE VERSUS INTIMIDATING IMPACT OF PITCH

Harmonic sounds are usually used by many species in affiliative contexts such as mating and other positive social interactions that often lead to increased social cohesion. Many species of birds, for instance, sing to advertise sexually,^{213,214} to maintain social bonds,²¹⁵ to coordinate behavior, and to facilitate complex social interactions.²¹⁶ Also, many species of mammals,²¹⁷ including primates,^{218,219} produce harmonic sounds in similar contexts. However, each vocal message in addition to affecting the selected recipient is exposed to eavesdropping. This threat concerns both the risk of being noticed by a predator^{220,221,132} and by a conspecific²²² competing for various resources with the sender of the message. As far as the risk of predation is concerned, a blending effect caused by pitch matching can make it difficult to recognize or trace a particular individual by a predator. At the same time amplitude summation of sounds produced during chorusing can contribute to the perceived body size exaggeration effect. In social species, if the risk of eavesdropping by conspecifics is high, natural selection can favor such affiliative signals which additionally contain a warning–threat message that encodes information about coalition strength or group cohesion, for example. In many cases of warning signals, including those that operate over long distances, it is the rhythm not the melody that seems to be primarily related to the deterrent function.²²³ A notable example of defensive signals where rhythm plays a crucial role is barking, which can be detected from long distances. In addition to being composed of rhythmic, noisy, and repetitive abrupt vocalizations, barking also often includes harmonic sounds.²²⁴ Although barking is usually associated with domestic dogs similar mobbing calls have been observed in many other taxa, such as rodents,²²⁵ birds,²²⁶ and primates.²²⁷ In fact, the exposure of primates to danger such as predators or intruders often involves repetitious and rhythmic calls,²²⁸ often similar to barking.^{163,227,229}

In many social species, including primates, rhythmic calls serve as long-distance territorial signals. For instance, it has been observed that white-handed gibbons (*Hylobates lar*) use songs to deter intrud-

ing conspecifics.^{230,189} Although auditory–motor synchronization, a key component of human musicality, has so far been observed in chimpanzees only in a rudimentary form,^{231,232} the vocalizations of primates such as lemurs,²³³ gibbons,^{234,235} and orangutans²³⁶ are characterized by isochrony or other types of categorical rhythms. Rhythm is also a particularly important feature of chimpanzee drumming,²³⁷ which is thought to be a specific type of agonistic display.^{238,239} Importantly, chimpanzee drumming is often accompanied by vocalizations that are also isochronous.²³⁷ All these observations suggest that the basic cognitive mechanisms that underlie the rhythmic part of human musicality were shared by our last common ancestor with chimpanzees. Moreover, the use of coordinated drumming by chimpanzees to intimidate conspecifics indicates that rhythm was probably a crucial component of hominin defensive signaling. For Hagen and Bryant the use of rhythm by hominins as a part of coalition signaling display had a precursor in such rhythmic territorial defense signals.²⁰ Also Fitch and Zuberbühler have proposed that defensive displays against predators could have been an adaptive function for hominin rhythmic vocal calls.²² They have emphasized that the effectiveness of synchronized group signaling is due to the acoustic summation of amplitudes, which increases signal loudness.

As synchrony necessitates coordination, it is a good marker of coalition strength which can both intimidate enemies and encourage cooperation.²⁰ Nevertheless, if the same signal is to fulfill two different functions, depending on the “friend or foe” perspective, pitch can also act as a ritualized tool of intimidation, especially if it is well matched and additionally well synchronized. In this context, the hypothesis that the human ability to perceive and synchronize to a beat may have evolved as a byproduct of vocal production learning,²⁴⁰ can be reversed. Synchronized displays may have emerged among Australopithecines—or even earlier—as a mechanism to deter predators²² or to signal dominance and cohesion among conspecifics.²⁰

Given this evolutionary context, it is plausible that the capacity to perceive and synchronize to a beat functioned as a preadaptation for vocal production learning in hominins. An important factor that may have additionally influenced the evolution of the ability to perceive and synchronize to a beat was the transition to bipedalism^{31,241} and endurance locomotion.²⁴² Given that human vocal production learning is a complex ability composed of multiple elements^{243–246} that likely evolved gradually over time, the imitation of the temporal order of sounds alone might have been one of the earliest components of vocal production learning in hominins. The fundamental role of rhythm in defense signaling suggests that the evolution of vocal learning in our lineage may have started exactly with the temporal control of vocalizations in the predecessors of the *Homo* genus, probably even in *Ardipithecus ramidus*.^{241,247} This development ultimately enhanced the vocal control of F_0 in *Homo erectus* and their descendants (see Figure 3).²⁴¹ Apart from this, synchronization can be viewed as a rhythm analogue of pitch matching as synchronization of sounds is in fact a kind of alignment and can act as a signal of cohesion.²⁰ This simplified form of alignment, compared to pitch matching, might have been inadequate in deterring conspecific intruders if the arms race among hominins resulted in the widespread adoption of this tool as a

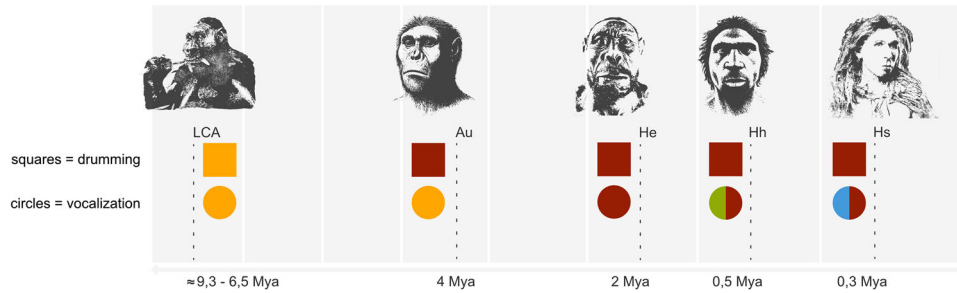


FIGURE 3 The evolution of vocal defense signaling in the *Homo sapiens* lineage: Yellow, isochrony/asynchrony; red, synchrony; green, heterophony; blue, pitch matching. Au, Australopithecines; He, *Homo erectus*; Hh, *Homo heidelbergensis*; Hs, *Homo sapiens*; LCA, last common ancestor of chimpanzees and humans.

defense strategy. After all, synchronization, being time-dependent,²⁴⁸ is inherently limited in complexity. Incorporating pitch matching into a synchronized display adds an extra dimension to signaling, thereby enriching the information conveyed.

In many cultures such displays are accompanied by rhythmic entrainment,²⁴⁹ which is not only an analogous form of pitch matching, but also appears to be a key mechanism for sustaining prolonged performance. In this context, the duration of the performance may serve as a significant indicator of stamina, further contributing to deterring conspecific intruders. From this perspective, rhythm is not a defining feature of short-distance defense signals but rather one of the signaling tools that can be enhanced by pitch. Specifically, rhythmic sequences of broadband sounds are less effective in propagation compared to rhythmic harmonic sounds, such as the duets of siamang²⁵⁰ and tarsier.²⁵¹ In humans, the combination of rhythm synchronization and pitch matching results in sequence alignment, which appears to be a powerful tool for exerting social influence over long distances.

Nowadays, musical behavior is used in many rituals in which the meaning is often related to combat, intimidation, and threat display,²⁵² as in the case of the Māori *haka* performance,²⁵³ football fan chants and songs,²⁵⁴ and in military songs.²⁵⁵⁻²⁵⁷ Lehmann et al. have indicated 20 ethnological examples of dual songs such as Inuit *piseq* songs, *tiv* songs from Nigeria, and musical territorial negotiations by Yanomami people from Venezuela.²⁵⁸ While musical duels represent a kind of competition between soloists, the use of ritualized, musicalized speech by competitors may reflect a need to emphasize the verbal message via an additional intimidating signal. Although, according to Lehmann et al., the ubiquity of this tradition can be a result of intrasexual selection, the presence of similar intimidating functions in the use of military and football fan songs may suggest that this function of music could have evolved due to selection by hominin social environments. Both in military songs and football fan chants, the message encoded by words seems to be less important than the influence of musical structure. This suggests that the intimidating power of music has its roots in pre-verbal collective rituals rather than being a product of intrasexual selection.

THE ROLE OF TIMBRE, PITCH, HETEROPHONY, AND VIBRATO IN SIGNALING GROUP SIZE

The size of a group plays a crucial role in determining its defensive effectiveness.²⁵⁹ Before launching an attack predators and conspecific intruders typically assess their likelihood of success, and a large group of defenders can serve as a formidable barrier—or even pose a counter-threat. As a result, accurately estimating group size is essential for intruders, while defenders may benefit from exaggerating their numbers. Additionally, the ability to signal group size over long distances can be advantageous as it may deter potential intruders from approaching the territory. In order to distinguish between different individuals by their voices, each voice must be processed in the auditory system as a separate auditory stream.²⁶⁰ Timbre and F_0 are typically effective cues for auditory stream segregation.²⁶¹⁻²⁶³ However, when simultaneous voices are similar in F_0 and timbre, separating them into distinct auditory streams becomes challenging, leading to the loss of information about the exact number of individuals. A potential solution to this issue may have been the use of varied timbres and pitches by individual hominins. This strategy bears resemblance to the chorusing behavior of howler monkeys whose loud roars are often used to repel opponents from rival groups.²⁶⁴ The distance over which these vocalizations can be heard often matches or even exceeds that of chimpanzee pant-hoots.¹⁶³ Nevertheless, in such cases, the range of audibility and recognition of the message, especially roars and barks,²⁶⁵ is limited²⁶⁶ compared to the solo barks or duets of gibbons²⁶⁷ or the group singing of humans in unison or octaves.

These acoustic constraints seem to require a trade-off between signaling group size and signal range. When a group consists of around a dozen individuals, its size can be estimated through auditory stream segregation. Certain musical features—such as varied timbres, vibrato, polyphony, and heterophony—undoubtedly facilitate this process. The idea that the evolutionary precursor to music was based on heterophony characterized by poor synchronization²⁶⁸ and pitch matching^{203,269} is consistent with this function. However, since the counting abilities of mammalian predators and hominin intruders were limited,²⁷⁰ in cases where a hominin group consisted of several dozen individuals, the estimation of group size likely relied upon approxima-

tion rather than precise counting of individuals. In such a case, it seems to be a better solution to signal the group size using sound intensity. As synchronized and pitch-matched signals increase sound intensity through amplitude summation, group singing in unison or in octaves may have served as an honest signal not only of group size but also of cooperativity.²⁷

In this respect, an important point not to be overlooked is the collective singing of men and women who can match pitches by singing in octaves. In fact, even if male and female voices differ in timbre those that share the same pitch are often difficult to distinguish.¹⁰⁴ However, the vocal intensity of individual singers may vary depending on their body size and vocal capacity, which makes it more difficult to accurately estimate group size. This could have opened the door to deceptive strategies, which may have in turn created selective pressure for enhanced abilities in synchronization, pitch matching, and vocal intensity. In this respect, group singing—unlike other forms of human vocalization—can operate on a much larger scale and bring together dozens of individuals.¹²⁷ Consequently, as group size increased, the tendency to diversify individual vocalizations in terms of timbre and pitch may have given way to selective pressure for pitch matching. This growth in group size may have also contributed to the evolution of octave equivalence, as the blending of male and female voices amplifies loudness, effectively signaling the group's size and deterring conspecific intruders by emphasizing group unity. It is also possible that the harmonic clarity that distinguishes *H. sapiens* vocalizations from those of many other primates¹⁰⁴ is a result of this evolutionary process.

VOCAL PRODUCTION LEARNING AS A PRECONDITION FOR MUSICAL PITCH

As the musical use of pitch in hominin history probably started with vocalizations,^{102,271} one can assume that the evolution of the pitch part of human musicality must have been related to singing. To sing culturally variable songs, the voluntary laryngeal control of F_0 had to evolve.²⁷¹ This capacity is a part of the broader ability called vocal production learning, which distinguishes *H. sapiens* from other living primates^{272,273} including chimpanzees, meaning that it had to evolve in our evolutionary lineage after the split from our last common ancestor with chimpanzees. However, it has been observed in some species of birds²⁷⁴ and mammals.^{272,245} Interestingly, some of these vocal learning species such as the red-winged blackbird (*Agelaius phoeniceus*),²⁷⁵ the yellow-breasted boubous (*Laniarius atrof lavus*),²⁷⁶ the banded wren (*Thryophilus pleurostictus*),¹⁶⁴ and the great Himalayan leaf-nosed bat (*Hipposideros armiger*)^{176,177} use their vocalizations as territorial defense signals, which implies that apart from sexual display and information sharing,²⁷⁷ one should also take into account a defensive function as a potential reason for the evolution of vocal learning among hominins. Moreover, considering the flexibility of use and developmental plasticity of primate calls,^{95,96} the gradual evolution of vocal learning may have even begun in the last common ancestors of humans and gibbons. It is worth emphasizing that developmental learning among primates continues into adulthood,²⁷⁸ influencing the plastic-

ity of vocal development in primates²⁷⁹ including chimpanzees.²⁸⁰ Although in general vocal production learning is defined as an ability to reproduce by voice what is heard,²⁸¹ vocal learners differ in their capability to reproduce sounds, meaning that vocal learning can be seen as a continuum of abilities rather than a yes or no ability.²⁴³ Moreover, these differences could concern various parameters such as the accuracy of the copy, the degree of possible modifications, and what is being copied.²⁸² It can be assumed that different adaptive functions that result in structurally different vocalizations should impose separate and specific characteristics on species-specific vocal learning. This difference is evident in the case of human vocally learned forms of communication such as speech and song.

As the role of the temporal envelope and temporal fine structure differs in the perception of speech and music^{283–286} one can suspect that vocal production learning of these two communicative forms is also based, at least partly, on the imitation of different acoustic features. This difference seems to be evident at the cognitive level since the recognition of musical structure is grounded in pitch and rhythm,^{4,287,288} whereas speech is perceived mainly as a chain of changing vowels and consonants that are recognizable due to the contrast of formants.²⁸⁹ Even the rhythms of speech and music, encoded through modulations of sound intensity over time differ and exhibit well-separated peaks—around 5 Hz for speech and 2 Hz for music.²⁹⁰ Consequently, the imitation of these two vocal expressions likely depends on distinct rhythmic biases. Since music differs from speech in its much higher degree of repetition,²⁹¹ recognizing these repetitions may also play an important role in distinguishing music from speech. The use of repetitive musical phrases, though likely rooted in more general abilities to recognize similar patterns, also requires vocal production learning of pitch. The crucial part of human vocal production learning that determines the use of pitch in music is the volitional control of F_0 . In fact, it is this ability that allows us to imitate the pitches we hear.²⁹² While pitch is also an important part of speech and can imply that the volitional control of F_0 evolved as a part of the language faculty,²⁹³ both the stability of F_0 ^{8,7} and its syntactic function²⁹⁴ seem to be much more important in music than in speech. Thus, the ability to sustain F_0 as a part of vocal production learning must have been selected for because of the adaptive functions that formed the basis of musicality and not the speech faculty. If the stability of F_0 was such an important acoustic feature of hominin protomusic, then it would have also become an indicator for music recognition and comprehension. This means that the volitional control of F_0 must have coevolved with a perceptive sensitivity to pitch stability (see Figure 4).

PITCH DISCRETENESS, COMBINATORIALITY, AND LONG-DISTANCE CALLS

The recognition and vocalization of sounds with different stable F_0 must have also become the basis for another important feature of musical pitch discreteness (see Figure 4). Only hominins endowed with these two abilities could have vocalized pitches in a discrete-like fashion—also in relation to discrete units of time—that characterize



FIGURE 4 Music-specific abilities related to pitch. Vocal production learning of F_0 , allows hominins to volitionally control and sustain F_0 and align their vocalized pitch with other singers; Sensitivity to F_0 stability, enables hominins to interpret harmonic sounds in terms of discrete pitch intervals; Sensitivity to tuning, enables the recognition of a culture-specific pitch system; Octave equivalence, restricts the scope of a pitch system; Relative pitch, enables hominins to recognize transposed melodies as the same pitch patterns; Sensitivity to melody, facilitates the veridical memory of vocalized melodies; Sensitivity to pitch distribution, enables the interpretation of pitch sequences in terms of tonal hierarchy.

contemporary singing and music.⁴ In other words, the emergence of pitch discreteness was related to the natural selection of the volitional control and sustenance of F_0 as well as sensitivity to pitch stability. This is not to suggest that the ability to produce a stable F_0 is an exceptional skill exclusive to humans. However, the difference between the stability of harmonic sounds in speech and in music may reflect the difference between the ranges of information transfer achieved by these two forms of communication. While speech is used to exchange information over a short distance, singing allows a signal to be transmitted over a longer distance. Since reverberations of soundwaves impede the intelligibility of calls characterized by rapid frequency modulations,¹⁵⁶ the use of stable F_0 can be one of the factors that facilitate long-distance communication, especially in a forested environment that causes many soundwave reverberations. This means that in such an environment, in order to send a message efficiently over a long distance it is better to code information in an acoustic signal by means of frequency rather than amplitude modulation. Although *Homo* species have been thought to live in open areas,²⁹⁵ some studies suggest that their habitats were much more diverse^{296,297} including forested landscapes.²⁹⁸ Thus, the ability to vocally control F_0 stability in singing could have evolved as an adaptation to long-distance vocal signaling.

Musical pitch is also important for the combinatoriality of music (see Figure 4).^{299,300} Music is claimed to be an example of the Humboldt system⁴ (i.e., a system which is based on a restricted number of discrete units combined according to certain rules). The big advantage of such a system is its power to generate an enormous number of expressions from a finite number of units.³⁰¹ An important feature of our

experience of harmonic sounds that restricts the number of pitches (units) is octave equivalence (see Figure 4).⁵ Due to octave equivalence, the whole range of audible sound frequencies is divided into equivalent parts—octaves which create a circularity of pitch space.^{5,302} Since octave-based pitch space appears to dominate musical cultures,¹⁷ and the ability to perceive octave equivalence seems independent of speech abilities³⁰³—despite influencing speech³⁰⁴—it should be considered a fundamental aspect of the pitch component of musicality. That is not to say that octave equivalence is specific solely to music and that it evolved because of selective pressures related to musicality. Since it has been observed in nonhuman species,^{114,115} the use of this ability in musical behavior should rather be considered an effect of repurposing a pre-existing cognitive tool.^{305,306} Independent of whether the ubiquity of octave equivalence in musical cultures is an effect of convergent cultural evolution,³⁰⁷ or a human-specific trait based on a perception-production feedback loop,¹²⁷ the widespread presence of octave in musical pitch systems indicates its importance for musicality. Although some scholars have questioned the universality of octave equivalence,³⁰⁸ relative pitch—the ability to recognize two-pitch patterns as the same melody if their F_0 follow logarithmic relationships (i.e., based on the same frequency ratios)—appears to be independent of cultural background.³⁰⁹ Thanks to relative pitch the identity of transposed melodies is preserved.

SENSITIVITY TO TUNING, PITCH MATCHING, AND THE TWO TYPES OF MEMORIES FOR PITCHES

An important consequence of pitch discreteness and circularity is the emergence of a musical pitch system. While the discreteness of musical pitch is a universal feature of music,^{10,310,17,311} musical pitch systems vary cross-culturally both in terms of the number of pitch classes and the size of pitch intervals.³¹² This means that musical pitch systems are culture-specific. Nevertheless, sensitivity to any musical pitch system starts very early in ontogenesis,^{313,314} which suggests that sensitivity to tuning is based on so-called experience-expectant plasticity.³¹⁵ This type of plasticity implies a genetically constrained developmental path¹¹⁶ suggesting that sensitivity to a musical pitch system could have also been selected as a part of our musicality.

However, although humans are quite good at the discrimination of even small differences in the F_0 of harmonic sounds,³¹⁶ even in a musical context³¹⁷ the recognition of pitch intervals allows for greater or lesser tolerance.^{3,318,319} Admittedly, the tuning of pitch intervals by singing individuals is usually less precise³²⁰ than one would expect from the human ability of F_0 discrimination. Nevertheless, the observed tendency of singing individuals toward intonation accuracy³²¹ indicates that sensitivity to tuning is not only a passive, but also an active part of human musicality (see Figure 4). Therefore, the consequence of active sensitivity to tuning is the ability to match pitches among individuals, not only the production of sounds in general but the selected spectral characteristics of these sounds such as F_0 and other harmonics.

The effect of collective singing based on the sensitivity to tuning can be called pitch matching among singing individuals. Aligning pitches not only imposes additional demands on working memory for pitch but also entails the special role of predictions. After all, to sing a particular melody together, singers need to know the pitches of which it consists. Contemporary humans are assumed to be endowed with two types of musical expectations—schematic and veridical.³²² While schematic expectations in the domain of pitch require sensitivity to pitch distribution, veridical expectations demand sensitivity to specific melodic patterns.

Pitch distribution has been shown to be one of the main foundations for pitch syntax—tonality^{12,323,324}—which has often been claimed to be an additional dimension of musical pitch (see Figure 4).³²⁵ This dimension, which is described as a hierarchy of pitches, is based on subtle sensations—tonal qualia¹²—that accompany the experience of musical pitch in tonal music. For some researchers, this hierarchy is merely a consequence of the harmonic structure of periodic sounds, as they explore, for example, the origins of pitch centers in functional consonance, which arises from the simple relationships between harmonics of particular pitches.³²⁶ In contrast to acoustic explanations, many cognitive mechanisms such as the exposure effect,³²³ prediction effect,¹² and physiological nonlinearities in the auditory pathway^{327,328} have been proposed to explain tonal hierarchy. All these proposals refer to the general cognitive mechanisms which may suggest that tonality is not based on music-specific abilities. From this perspective, the emergence of tonality resulted from a solely cultural transmission process.¹⁴⁶

However, from a behavioral point of view, tonality (in the sense of a psychological hierarchy of pitches based on tonal center) is a specific psychological trait unique solely to music, which is a challenge for any explanation that refers to general cognitive mechanisms. One can imagine that purely cultural intergenerational transmission of musical patterns was enough for the emergence of tonality by means of a learner bottleneck.³²⁹ According to this explanation, tonal hierarchy should have become part of musical culture without any external cause or function needed to do so. In this case, sensitivity to pitch distribution is based on another type of plasticity—experience-dependent plasticity.³¹⁵ A possible alternative to the purely cultural origin of tonality is the canalization of a specific integration of many cognitive mechanisms that arose as a phenotypic adaptation in response to environmental challenges.^{32,306} As a result of this, tonality is a widespread feature of world music¹¹ and is acquired implicitly.^{330,331} Thus, it is probable that the sensitivity to pitch distribution is also a music-specific kind of the experience—expectant plasticity, and tonality is therefore evolutionary in origin.³² Although older adults seem to refine their schematic expectations together with their lifelong musical experience, both younger and older adults are equally effective at the recognition of unexpected pitches,³³² which also speaks for this explanation.

In contrast, the memory of melodic patterns seems to improve with the increasing experience of music and age,³³³ suggesting its reliance on experience-dependent plasticity.³¹⁵ This kind of developmental plasticity is lifelong among humans and allows us to learn completely

new invented activities for which we did not evolve.¹¹⁶ This is not to say that sensitivity to melody is based solely on a general-domain mechanism. Quite the opposite; the fact that people memorize sung melodies better than melodies played on instruments³³⁴ and that this advantage is independent of other biologically important signals coded in singers' timbre such as attractiveness,³³⁵ may suggest the evolutionary origin of vocalized pitch sequences as being part of the hominin communicative repertoire. However, the cultural variability of pitch sequences observed in contemporary musical cultures indicates the crucial role of vocal culture in the shaping of our sensitivity to melody (see Figure 4).

CONCLUSION

As it has been presented, the suggested function of musical pitch as defensive signaling could have coexisted with other functions such as social bonding,^{32,24,18,19,336} signaling coalition,^{34,44,20} and free rider recognition.³³ Rather than being mutually exclusive, they likely reinforced each other and created a synergistic loop that accelerated the evolution of musicality. One can argue, however, that hominins could have achieved the same functions without any of the abilities mentioned above. For instance, Merker raised the question of whether humans need music to bond at all, pointing out that nonhuman animals perform this task very well without music.³³⁷ The same logic can be used to question intraspecific defense signaling as an adaptive function of musical pitch. However, the hominin environment was characterized by increasing multilevel social complexity, which not only became a new specific source of selection, but also involved the creation of new niches with new challenges. Under such circumstances, adaptations that were effective in an environment of lower social complexity may not have been sufficient when hominin social complexity reached higher levels. Factors such as in-group competition for resources, division of labor in a group, the level of cooperation, and limits of social interactions may have influenced the evolution of cognitive capabilities.³³⁸ For example, it has been indicated that the increase in social complexity could be one of the reasons for the selection of more complex communicative abilities.^{339,76} Social complexity has also been claimed to be a source of the selection for abilities to sustain social cohesion.³⁴⁰ It is, therefore, reasonable to assume that in the case of a larger number of group members together with the changing aforesaid social factors, functions such as social bonding, signaling coalition or commitment, free rider recognition, as well as deterring foes could have also required new, more elaborate methods. The use of musical pitch as a defense signal is probably a type of an elaboration of more general signaling by means of rhythm synchronization.

The hypothesis that musical pitch may serve as a defensive function provides a new perspective on the evolution of musical behavior, integrating this function as a potential contributing factor without excluding other influences on human musicality. Importantly, this hypothesis is empirically testable through both observation and experimentation. As discussed, music continues to be employed in defensive contexts today and comparing pitch characteristics in such music with

those in noncombat-related music can clarify its specific defensive role. Additionally, human physiological and behavioral responses to variations in pitch can be measured, offering further insight into its adaptive significance. While further research is necessary to fully assess the evolutionary implications of musical pitch, this hypothesis expands the scope of inquiry and highlights previously unexplored pathways in the study of musicality.

ACKNOWLEDGMENTS

This research was funded in whole by National Science Centre, Poland [Grant number 2021/41/B/HS1/00541]. I would like to thank the reviewers for their immense effort and inspiring suggestions, which greatly enhanced my article. I am also grateful to Joseph Jordania for encouraging me to explore the issue of defense signaling and for his valuable comments on this text. My appreciation extends to Liliana Bether, whose input added aesthetic refinement to my diagrams, and to Peter Kośmider-Jones for his linguistic consultations.

CONFLICT OF INTEREST STATEMENT

The author declares no conflicts of interest.

ORCID

Piotr Podlipniak  <https://orcid.org/0000-0002-4326-559X>

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/nyas.70000>.

REFERENCES

1. Song, X., Osmanski, M. S., Guo, Y., & Wang, X. (2016). Complex pitch perception mechanisms are shared by humans and a New World monkey. *Proceedings of the National Academy of Sciences*, 113, 781–786. <https://doi.org/10.1073/pnas.1516120113>
2. Norman-Haignere, S. V., Kanwisher, N., McDermott, J. H., & Conway, B. R. (2019). Divergence in the functional organization of human and macaque auditory cortex revealed by fMRI responses to harmonic tones. *Nature Neuroscience*, 22, 1057–1060. <https://doi.org/10.1038/s41593-019-0410-7>
3. Stainsby, T., & Cross, I. (2008). The perception of pitch. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 47–58). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199298457.013.0005>
4. Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae*, 6, 3–21. <https://doi.org/10.1177/102986490200600101>
5. Deutsch, D., Dooley, K., & Henthorn, T. (2008). Pitch circularity from tones comprising full harmonic series. *Journal of the Acoustical Society of America*, 124, 589–597. <https://doi.org/10.1121/1.2931957>
6. Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *Journal of the Acoustical Society of America*, 129, 2245–2252. <https://doi.org/10.1121/1.3562174>
7. Ozaki, Y., Tierney, A., Pfordresher, P. Q., McBride, J. M., Benetos, E., Proutskova, P., Chiba, G., Liu, F., Jacoby, N., Purdy, S. C., Opondo, P., Fitch, W. T., Hegde, S., Rocamora, M., Thorne, R., Nweke, F., Sadaphal, D. P., Sadaphal, P. M., Hadavi, S., ... Savage, P. E. (2024). Globally, songs and instrumental melodies are slower and higher and use more stable pitches than speech: A registered report. *Science Advances*, 10, eadm9797. <https://doi.org/10.1126/sciadv.adm9797>
8. Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, 10, Article e1001372. <https://doi.org/10.1371/journal.pbio.1001372>
9. Simchy-Gross, R., & Margulis, E. H. (2018). The sound-to-music illusion. *Music & Science (London)*, 1, 1–6. <https://doi.org/10.1177/2059204317731992>
10. Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *PNAS*, 112, 8987–8992. <https://doi.org/10.1073/pnas.1414495112>
11. Mehr, S. A., Singh, M., Knox, D., Ketter, D., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science* (1979), 366, eaax0868. <https://doi.org/10.1126/science.aax0868>
12. Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. The MIT Press.
13. Pinker, S. (1997). *How the mind works*. Norton.
14. Trainor, L. J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 370, 20140089. <https://doi.org/10.1098/rstb.2014.0089>
15. Purves, D. (2017). *Music as biology: The tones we like and why*. Harvard University Press.
16. Savage, P. E., Passmore, S., Chiba, G., Currie, T. E., Suzuki, H., & Atkinson, Q. D. (2022). Sequence alignment of folk song melodies reveals cross-cultural regularities of musical evolution. *Current Biology*, 32, 1395–1402.e8. <https://doi.org/10.1016/j.cub.2022.01.039>
17. Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, 41, 229–248. <https://doi.org/10.1177/0305735611425896>
18. Roederer, J. G. (1984). The search for a survival value of music. *Music Perception: An Interdisciplinary Journal*, 1, 350–356. <https://doi.org/10.2307/40285265>
19. Anthony, S. (1992). *Music and the mind*. Ballantine Books.
20. Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, 14, 21–51. <https://doi.org/10.1007/s12110-003-1015-z>
21. Hagen, E. H. (2022). The biological roots of music and dance. *Human Nature*, 33, 261–279. <https://doi.org/10.1007/s12110-022-09429-9>
22. Fitch, W. T., & Zuberbühler, K. (2024). Group vocal displays, anti-predator behaviour and the evolution of musicality. In J. Jordania & J. Wade (Eds.), *Defense strategies in early human evolution* (pp. 15–34). Logos.
23. Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagen, E. H. (2021). Origins of music in credible signaling. *Behavioral and Brain Sciences*, 44, e60–e60. <https://doi.org/10.1017/S0140525x20000345>
24. Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 44, e59:1–e59:22. <https://doi.org/10.1017/S0140525x20000333>
25. Darwin, C. (1871). *The descent of man, and selection in relation to sex* (1st ed.). John Murray.
26. Miller, G. F. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 329–360). The MIT Press. <https://doi.org/10.1177/004057368303900411>
27. Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 315–327). The MIT Press.
28. Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 103–123). The MIT Press.

29. Ravnani, A. (2018). Darwin, sexual selection, and the origins of music. *Trends in Ecology & Evolution*, 33, 716–719. <https://doi.org/10.1016/j.tree.2018.07.006>
30. Podlipniak, P. (2024). The role of pitch in choral singing as an intraspecific defense strategy. In J. Jordania (Ed.), *Defense strategies in early human evolution* (pp. 50–57). Logos.
31. Mithen, S. J. (2009). The music instinct: The evolutionary basis of musicality. *Annals of the New York Academy of Sciences*, 1169, 3–12. <https://doi.org/10.1111/j.1749-6632.2009.04590.x>
32. Podlipniak, P. (2016). The evolutionary origin of pitch centre recognition. *Psychology of Music*, 44, 527–543. <https://doi.org/10.1177/0305735615577249>
33. Podlipniak, P. (2023). Free rider recognition—A missing link in the Baldwinian model of music evolution. *Psychology of Music*, 51, 1397–1413. <https://doi.org/10.1177/03057356221129319>
34. Hagen, E. H., & 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. *Musicae Scientiae*, 13, 291–320. <https://doi.org/10.1177/1029864909013002131>
35. Honing, H. (2018). On the biological basis of musicality. *Annals of the New York Academy of Sciences*, 1423, 51–56. <https://doi.org/10.1111/nyas.13638>
36. Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370, 20140091. <https://doi.org/10.1098/rstb.2014.0091>
37. Brown, S. (2022). Group dancing as the evolutionary origin of rhythmic entrainment in humans. *New Ideas in Psychology*, 64, 100902. <https://doi.org/10.1016/j.newideapsych.2021.100902>
38. Ravnani, A. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology*, 5, 1118.
39. Kotz, S. A., Ravnani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, 22, 896–910. <https://doi.org/10.1016/j.tics.2018.08.002>
40. McNeill, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Harvard University Press.
41. Mithen, S. J. (2006). *The singing Neanderthals: The origins of music, language, mind, and body*. Harvard University Press.
42. Dunbar, R. I. M. (2012). On the evolutionary function of song and dance. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 201–214). Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199227341.003.0008>
43. Mehr, S. A., & Krasnow, M. M. (2017). Parent-offspring conflict and the evolution of infant-directed song. *Evolution and Human Behavior*, 38, 674–684. <https://doi.org/10.1016/j.evolhumbehav.2016.12.005>
44. Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagen, E. H. (2021). Origins of music in credible signaling. *Behavioral and Brain Sciences*, 44, e60. <https://doi.org/10.1017/S0140525x20000345>
45. Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, 16, 285–295. [https://doi.org/10.1016/0163-6383\(93\)80036-8](https://doi.org/10.1016/0163-6383(93)80036-8)
46. Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, 20, 383–396. [https://doi.org/10.1016/S0163-6383\(97\)90009-6](https://doi.org/10.1016/S0163-6383(97)90009-6)
47. Trainor, L. J. (1996). Infant preferences for infant-directed versus noninfant-directed playsongs and lullabies. *Infant Behavior and Development*, 19, 83–92. [https://doi.org/10.1016/S0163-6383\(96\)90046-6](https://doi.org/10.1016/S0163-6383(96)90046-6)
48. Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., ... Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, 6, 1545–1556. <https://doi.org/10.1038/s41562-022-01410-x>
49. Juslin, P. N. (2021). Mind the gap: The mediating role of emotion mechanisms in social bonding through musical activities. *Behavioral and Brain Sciences*, 44, e80. <https://doi.org/10.1017/S0140525x2000120X>
50. Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Toward inclusive theories of the evolution of musicality. *Behavioral and Brain Sciences*, 44, e121. <https://doi.org/10.1017/S0140525x21000042>
51. Jordania, J. (2008). Origins of rhythm and the defense strategy of human ancestors. In *Problems of traditional polyphony: Materials of the third international symposium on traditional polyphony* (pp. 47–65). Tbilisi V. Ssaradjishvili State Conservatoire.
52. Jordania, J. (2011). *Why do people sing? Music in human evolution*. Logos.
53. Knight, C., & Lewis, J. (2017). Wild voices: Mimicry, reversal, metaphor, and the emergence of language. *Current Anthropology*, 58, 435–453. <https://doi.org/10.1086/692905>
54. Jordania, J. (2024). Music as aposematic signal: Predator defense strategies in early human evolution. *Frontiers in Psychology*, 14, Article 1271854. <https://doi.org/10.3389/fpsyg.2023.1271854>
55. Jordania, J. (2024). Warning display as a strategy of defense and offense in early human evolution. In J. Jordania & J. Wade (Eds.), *Defense strategies in early human evolution* (pp. 58–83). Logos.
56. Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, 102, 1213–1222. <https://doi.org/10.1121/1.421048>
57. Anikin, A., Pisanski, K., & Reby, D. (2022). Static and dynamic formant scaling conveys body size and aggression. *Royal Society Open Science*, 9, 211496. <https://doi.org/10.1098/rsos.211496>
58. McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, 47, 379–387. <https://doi.org/10.1006/anbe.1994.1052>
59. Zanette, L. Y., Frizzelle, N. R., Clinchy, M., Peel, M. J. S., Keller, C. B., Huebner, S. E., & Packer, C. (2023). Fear of the human “super predator” pervades the South African savanna. *Current Biology*, 33, 4689–4696.e4. <https://doi.org/10.1016/j.cub.2023.08.089>
60. Rodriguez, J., Hölzchen, E., Caso-Alonso, A. I., Berndt, J. O., Hertler, C., Timm, I. J., & Mateos, A. (2023). Computer simulation of scavenging by hominins and giant hyenas in the late Early Pleistocene. *Scientific Reports*, 13, Article 14283. <https://doi.org/10.1038/s41598-023-39776-1>
61. Bunn, H. T., & Ezzo, J. A. (1993). Hunting and scavenging by Plio-Pleistocene hominids: Nutritional constraints, archaeological patterns, and behavioural implications. *Journal of Archaeological Science*, 20, 365–398. <https://doi.org/10.1006/jasc.1993.1023>
62. Pobiner, B. L. (2020). The zooarchaeology and paleoecology of early hominin scavenging. *Evolutionary Anthropology: Issues, News, and Reviews*, 29, 68–82. <https://doi.org/10.1002/evan.21824>
63. Parkinson, J. A. (2018). Revisiting the hunting-versus-scavenging debate at FLK Jinj: A GIS spatial analysis of bone surface modifications produced by hominins and carnivores in the FLK 22 assemblage, Olduvai Gorge, Tanzania. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 511, 29–51. <https://doi.org/10.1016/j.palaeo.2018.06.044>
64. Lehmann, K. D. S., Goldman, B. W., Dworkin, I., Bryson, D. M., & Wagner, A. P. (2014). From cues to signals: Evolution of interspecific communication via aposematism and mimicry in a predator-prey system. *PLoS ONE*, 9, e91783.
65. Laidre, M. E., & Johnstone, R. A. (2013). Animal signals. *Current Biology*, 23, R829–R833. <https://doi.org/10.1016/j.cub.2013.07.070>
66. Turner, C. R., Spike, M., & Magrath, R. D. (2023). The evolution of eavesdropping on heterospecific alarm calls: Relevance, reliability,

- and personal information. *Ecology & Evolution*, 13, Article e10272. <https://doi.org/10.1002/ece3.10272>
67. Trillo, P. A., Benson, C. S., Caldwell, M. S., Lam, T. L., Pickering, O. H., & Logue, D. M. (2019). The influence of signaling conspecific and heterospecific neighbors on eavesdropper pressure. *Frontiers in Ecology and Evolution*, 7, Article 292. <https://doi.org/10.3389/fevo.2019.00292>
 68. Magrath, R. D., Haff, T. M., Fallow, P. M., & Radford, A. N. (2015). Eavesdropping on heterospecific alarm calls: From mechanisms to consequences. *Biological Reviews*, 90, 560–586. <https://doi.org/10.1111/brv.12122>
 69. Ha, J., Lee, K., Yang, E., Kim, W., Song, H., Hwang, I., Lee-Cruz, L., Park, J., Song, J., Park, C.-R., Lee, W., Jablonski, P., & Lee, S.-I. (2023). Unlearned adaptive responses to heterospecific referential alarm calls in two bird species from separate evolutionary lineages. *Scientific Reports*, 13, Article 20287. <https://doi.org/10.1038/s41598-023-47052-5>
 70. Haff, T. M., & Magrath, R. D. (2012). Learning to listen? Nestling response to heterospecific alarm calls. *Animal Behaviour*, 84, 1401–1410. <https://doi.org/10.1016/j.anbehav.2012.09.005>
 71. Wheeler, B. C., Fahy, M., & Tiddi, B. (2019). Experimental evidence for heterospecific alarm signal recognition via associative learning in wild capuchin monkeys. *Animal Cognition*, 22, 687–695. <https://doi.org/10.1007/s10071-019-01264-3>
 72. Townsend, C., Ferraro, J. V., Habecker, H., & Flinn, M. V. (2023). Human cooperation and evolutionary transitions in individuality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378, 20210414. <https://doi.org/10.1098/rstb.2021.0414>
 73. Hart, D., & Sussman, R. W. (2011). The influence of predation on primate and early human evolution: Impetus for cooperation. In R. W. Sussman & C. R. Cloninger (Eds.), *Origins of altruism and cooperation* (pp. 19–40). Springer New York. https://doi.org/10.1007/978-1-4419-9520-9_3
 74. Kortlandt, A. (1980). How might early hominids have defended themselves against large predators and food competitors? *Journal of Human Evolution*, 9, 79–112. [https://doi.org/10.1016/0047-2484\(80\)90066-4](https://doi.org/10.1016/0047-2484(80)90066-4)
 75. Willems, E. P., & van Schaik, C. P. (2017). The social organization of *Homo ergaster*: Inferences from anti-predator responses in extant primates. *Journal of Human Evolution*, 109, 11–21. <https://doi.org/10.1016/j.jhevol.2017.05.003>
 76. Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1785–1801. <https://doi.org/10.1098/rstb.2011.0213>
 77. Freeberg, T. M. (2006). Social complexity can drive vocal complexity: Group size influences vocal information in Carolina chickadees. *Psychological Science*, 17, 557–561. <https://doi.org/10.1111/j.1467-9280.2006.01743.x>
 78. Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2017). Structural variability and communicative complexity in acoustic communication. *Animal Behaviour*, 134, 229–237. <https://doi.org/10.1016/j.anbehav.2016.06.012>
 79. Lomax, A. (1968). *Folk song style and culture: A staff report on cantometrics*. American Association for the Advancement.
 80. Blacking, J. (1973). *How musical is man?* University of Washington Press.
 81. Lomax, A. (1962). Song structure and social structure. *Ethnology*, 1, 425. <https://doi.org/10.2307/3772850>
 82. Moser, C., Ackerman, J., Dayer, A., Proksch, S., & Smaldino, P. E. (2021). Why don't cockatoos have war songs? *Behavioral and Brain Sciences*, 44, e108. <https://doi.org/10.1017/S0140525x20001223>
 83. van der Dennen, J. M. G. (1995). *The origin of war: The evolution of a male-coalitional reproductive strategy*. Origin Press.
 84. Belinchón, M., Polo, P., Rodríguez-Sickert, C., Figueroa, O., Valenzuela, N., Pavez, P., & Muñoz-Reyes, J. A. (2024). The effect of intergroup competition outcome on ingroup cooperation: Insights from the male warrior hypothesis. *Frontiers in Psychology*, 15, 1303372.
 85. Van Hooff, J. A. R. A. M. (1990). Intergroup competition and conflict in animals and man. In J. M. G. van der Dennen & V. S. E. Falger (Eds.), *Sociobiology and conflict: Evolutionary perspectives on competition, cooperation, violence and warfare* (pp. 23–54). Springer Netherlands. https://doi.org/10.1007/978-94-009-1830-6_2
 86. Wrangham, R. W., & Glowacki, L. (2012). Intergroup aggression in chimpanzees and war in Nomadic hunter-gatherers. *Human Nature*, 23, 5–29. <https://doi.org/10.1007/s12110-012-9132-1>
 87. Manson, J. H., Wrangham, R. W., Boone, J. L., Chapais, B., Dunbar, R. I. M., Ember, C. R., Irons, W., Marchant, L. F., McGrew, W. C., Nishida, T., Paterson, J. D., Smith, E. A., Stanford, C. B., & Worthman, C. M. (1991). Intergroup aggression in chimpanzees and humans [and comments and replies]. *Current Anthropology*, 32, 369–390.
 88. Fabre, V., Condeani, S., Degioanni, A., & Herrscher, E. (2011). Neanderthals versus modern humans: Evidence for resource competition from isotopic modelling. *International Journal of Evolutionary Biology*, 2011, 1–16. <https://doi.org/10.4061/2011/689315>
 89. Gilpin, W., Feldman, M. W., & Aoki, K. (2016). An ecocultural model predicts Neanderthal extinction through competition with modern humans. *Proceedings of the National Academy of Sciences*, 113, 2134–2139. <https://doi.org/10.1073/pnas.1524861113>
 90. Ashton, B. J., Kennedy, P., & Radford, A. N. (2020). Interactions with conspecific outsiders as drivers of cognitive evolution. *Nature Communications*, 11, 4937. <https://doi.org/10.1038/s41467-020-18780-3>
 91. Sterelny, K. (2007). Social intelligence, human intelligence and niche construction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 719–730. <https://doi.org/10.1098/rstb.2006.2006>
 92. Fischer, J., & Price, T. (2017). Meaning, intention, and inference in primate vocal communication. *Neuroscience and Biobehavioral Reviews*, 82, 22–31. <https://doi.org/10.1016/j.neubiorev.2016.10.014>
 93. Winter, P., Handley, P., Ploog, D., & Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour*, 47, 230–239. <https://doi.org/10.1163/156853973x00085>
 94. Seyfarth, R. M., & Cheney, D. (2018). Pragmatic flexibility in primate vocal production. *Current Opinion in Behavioral Sciences*, 21, 56–61. <https://doi.org/10.1016/j.cobeha.2018.02.005>
 95. Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, 115, 1974–1979. <https://doi.org/10.1073/pnas.1717572115>
 96. Adret, P. (2022). Developmental plasticity in primate coordinated song: Parallels and divergences with duetting songbirds. *Frontiers in Ecology and Evolution*, 10, 862196.
 97. Leal, M., & Losos, J. B. (2015). A naturalist's insight into the evolution of signal redundancy. *American Naturalist*, 186, ii–iv. <https://doi.org/10.1086/682704/ASSET/IMAGES/LARGE/FG2.JPEG>
 98. Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford University Press.
 99. Eleuteri, V., Henderson, M., Soldati, A., Badihi, G., Zuberbühler, K., & Hobaier, C. (2022). The form and function of chimpanzee buttress drumming. *Animal Behaviour*, 192, 189–205. <https://doi.org/10.1016/j.anbehav.2022.07.013>
 100. Hobaier, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, 71, 96. <https://doi.org/10.1007/s00265-017-2325-1>
 101. Taylor, D., Dezechache, G., & Davila-Ross, M. (2021). Filling in the gaps: Acoustic gradation increases in the vocal ontogeny of chimpanzees

- (*Pan troglodytes*). *American Journal of Primatology*, 83, e23249. <https://doi.org/10.1002/ajp.23249>
102. Morley, I. (2013). *The prehistory of music: Human evolution, archaeology, and the origins of musicality*. Oxford University Press.
 103. McAdams, S., & Bregman, A. (1979). Hearing musical streams. *Computer Music Journal*, 3, 26–43.
 104. Wagner, B., & Hoeschele, M. (2022). The links between pitch, timbre, musicality, and social bonding from cross-species research. *Comparative Cognition & Behavior Reviews*, 17, 13–32. <https://doi.org/10.3819/CCBR.2022.170002>
 105. Hoeschele, M. (2017). Animal pitch perception: Melodies and harmonies. *Comparative Cognition & Behavior Reviews*, 12, 5–18. <https://doi.org/10.3819/CCBR.2017.120002>
 106. Searcy, W. A., Soha, J., Peters, S., & Nowicki, S. (2021). Variation in vocal production learning across songbirds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376, 20200257. <https://doi.org/10.1098/rstb.2020.0257>
 107. Reiss, D., & McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): Evidence for vocal learning. *Journal of Comparative Psychology*, 107, 301–312. <https://doi.org/10.1037/0735-7036.107.3.301>
 108. Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, 55, 673–687. <https://doi.org/10.1006/anbe.1997.0602>
 109. Doolittle, E. L., Gingras, B., Endres, D. M., & Fitch, W. T. (2014). Overtone-based pitch selection in hermit thrush song: Unexpected convergence with scale construction in human music. *Proceedings of the National Academy of Sciences*, 111, 16616–16621. <https://doi.org/10.1073/PNAS.1406023111>
 110. Richner, H. (2016). Interval singing links to phenotypic quality in a songbird. *Proceedings of the National Academy of Sciences*, 113, 12763–12767. <https://doi.org/10.1073/pnas.1610062113>
 111. Doolittle, E., & Brumm, H. (2012). O Canto do Uirapuru: Consonant pitches and patterns in the song of the musician wren. *Journal of Interdisciplinary Music Studies*, 6, 55–85. <https://doi.org/10.4407/jims.2013.10.003>
 112. McPherson, M. J., Dolan, S. E., Durango, A., Ossandon, T., Valdés, J., Undurraga, E. A., Jacoby, N., Godoy, R. A., & McDermott, J. H. (2020). Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nature Communications*, 11, 2786. <https://doi.org/10.1038/s41467-020-16448-6>
 113. Demany, L., Monteiro, G., Semal, C., Shamma, S., & Carlyon, R. P. (2021). The perception of octave pitch affinity and harmonic fusion have a common origin. *Hearing Research*, 404, 108213. <https://doi.org/10.1016/j.heares.2021.108213>
 114. Richards, D. G., Wolz, J. P., & Herman, L. M. (1984). Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, *Tursiops truncatus*. *Journal of Comparative Psychology*, 98, 10–28. <https://doi.org/10.1037/0735-7036.98.1.10>
 115. Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiwirth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology*, 129, 291–307.
 116. Patel, A. D. (2023). Human musicality and gene-culture coevolution: Ten concepts to guide productive exploration. In E. H. Margulis, P. Loui, & D. Loughridge (Eds.), *The science-music borderlands: Reckoning with the past and imagining the future* (pp. 15–37). The MIT Press. <https://doi.org/10.7551/mitpress/14186.003.0006>
 117. Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, 3, 272–279. [https://doi.org/10.1016/S1364-6613\(99\)01333-9](https://doi.org/10.1016/S1364-6613(99)01333-9)
 118. Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, 40, 1–18. <https://doi.org/10.1159/000261678>
 119. Ohala, J. J. (1995). The frequency code underlies the sound-symbolic use of voice pitch. In J. Nichols, J. J. Ohala, & L. Hinton (Eds.), *Sound symbolism* (pp. 325–347). Cambridge University Press. <https://doi.org/10.1017/CBO9780511751806.022>
 120. Pisanski, K., Cartei, V., Mcgettigan, C., Raine, J., & Reby, D. (2016). Voice modulation: A window into the origins of human vocal control? *Trends in Cognitive Sciences*, 20, 304–318. <https://doi.org/10.1016/j.tics.2016.01.002>
 121. Bidelman, G. M. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Frontiers in Psychology*, 4, 264.
 122. Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6, 37–46.
 123. Oxenham, A. J. (2019). Pitch: Perception and neural coding. In *Foundations in music psychology: Theory and research* (pp. 3–32). The MIT Press.
 124. Oxenham, A. J. (2013). Revisiting place and temporal theories of pitch. *Acoustical Science and Technology*, 34, 388–396. <https://doi.org/10.1250/ast.34.388>
 125. Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, 33, 185–191. [https://doi.org/10.1016/S0896-6273\(01\)00580-3](https://doi.org/10.1016/S0896-6273(01)00580-3)
 126. Cerca, J. (2023). Understanding natural selection and similarity: Convergent, parallel and repeated evolution. *Molecular Ecology*, 32, 5451–5462. <https://doi.org/10.1111/mec.17132>
 127. Bannan, N., Dunbar, R., & Bamford, J. (2024). The evolution of gender dimorphism in the human voice: The role of octave equivalence. *Current Anthropology*, 65, 503–527.
 128. Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., Welling, L. L. M., Dawood, K., Cárdenas, R., Burriss, R. P., Jablonski, N. G., Shriver, M. D., Weiss, D., Lameira, A. R., Apicella, C. L., Owren, M. J., Barelli, C., Glenn, M. E., & Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20152830. <https://doi.org/10.1098/rspb.2015.2830>
 129. Chen, S., Han, C., Wang, S., Liu, X., Wang, B., Wei, R., & Lei, X. (2022). Hearing the physical condition: The relationship between sexually dimorphic vocal traits and underlying physiology. *Frontiers in Psychology*, 13, 983688.
 130. Carollo, A., Montefalcone, P., Bornstein, M. H., & Esposito, G. (2023). A scientometric review of infant cry and caregiver responsiveness: Literature trends and research gaps over 60 years of developmental study. *Children*, 10, 1042. <https://doi.org/10.3390/children10061042>
 131. Young, K. S., Parsons, C. E., Jegindoe Elmholt, E.-M., Woolrich, M. W., Van Hartevelt, T. J., Stevner, A. B. A., Stein, A., & Kringelbach, M. L. (2016). Evidence for a caregiving instinct: Rapid differentiation of infant from adult vocalizations using magnetoencephalography. *Cerebral Cortex*, 26, 1309–1321. <https://doi.org/10.1093/cercor/bhv306>
 132. Thévenet, J., Papet, L., Coureaud, G., Boyer, N., Levréro, F., Grimault, N., & Mathevon, N. (2023). Crocodile perception of distress in hominid baby cries. *Proceedings of the Royal Society B: Biological Sciences*, 290, 20230201. <https://doi.org/10.1098/rspb.2023.0201>
 133. Flack, J. C., & de Waal, F. (2007). Context modulates signal meaning in primate communication. *Proceedings of the National Academy of Sciences*, 104, 1581–1586. <https://doi.org/10.1073/pnas.0603565104>
 134. Ford, J. K. B. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Canadian Journal of Zoology*, 69, 1454–1483. <https://doi.org/10.1139/z91-206>
 135. Sayigh, L., Tyack, P., Wells, R., & Scott, M. (1990). Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: Stability and mother-offspring comparisons. *Behavioral Ecology and Sociobiology*, 26, 247–260. <https://doi.org/10.1007/BF00178318/METRICS>
 136. Feeles, F. (1982). Song mimesis within colonies of *Cacicus c. cela* (Icteridae, Aves): A colonial password? *Zeitschrift für Tierpsychologie*, 58, 119–152. <https://doi.org/10.1111/j.1439-0310.1982.tb00312.x>

137. Hauber, M. E., Russo, S. A., & Sherman, P. W. (2001). A password for species recognition in a brood-parasitic bird. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268, 1041–1048. <https://doi.org/10.1098/RSPB.2001.1617>
138. Colombelli-Négre, D., Hauber, M. E., Robertson, J., Sulloway, F. J., Hoi, H., Griggio, M., & Kleindorfer, S. (2012). Embryonic learning of vocal passwords in superb Fairy-Wrens reveals intruder cuckoo nestlings. *Current Biology*, 22, 2155–2160. <https://doi.org/10.1016/j.CUB.2012.09.025>
139. Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4, 258–267. [https://doi.org/10.1016/S1364-6613\(00\)01494-7](https://doi.org/10.1016/S1364-6613(00)01494-7)
140. Polyanskaya, L., Ordin, M., & Busa, M. G. (2016). Relative salience of speech rhythm and speech rate on perceived foreign accent in a second language. *Language and Speech*, 60, 333–355. <https://doi.org/10.1177/0023830916648720>
141. Leong, V., Stone, M. A., Turner, R. E., & Goswami, U. (2014). A role for amplitude modulation phase relationships in speech rhythm perception. *Journal of the Acoustical Society of America*, 136, 366–381. <https://doi.org/10.1121/1.4883366>
142. Meyer, J., Dentel, L., & Meunier, F. (2013). Speech recognition in natural background noise. *PLoS ONE*, 8, e79279.
143. Meyer, J. (2015). *Whistled languages: A worldwide inquiry on human whistled speech*. Springer-Verlag.
144. Marler, P. (2008). The instinct to learn. In M. H. Johnson, Y. Munakata, & R. O. Gilmore (Eds.), *Brain development and cognition* (2nd ed., pp. 305–329). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470753507.CH17>
145. Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication*. (2nd ed.). Sinauer Associates.
146. Merker, B., Morley, I., & Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20140095. <https://doi.org/10.1098/rstb.2014.0095>
147. Tavernier, C., Ahmed, S., Hout, K. A., & Yeon, S. C. (2020). Feline vocal communication. *Journal of Veterinary Science*, 21, e18. <https://doi.org/10.4142/jvs.2020.21.e18>
148. Déaux, É. C., & Clarke, J. A. (2013). Dingo (*Canis lupus dingo*) acoustic repertoire: Form and contexts. *Behaviour*, 150, 75–101. <https://doi.org/10.1163/1568539X-00003038>
149. Faragó, T., Pongrácz, P., Range, F., Virányi, Z., & Miklósi, Á. (2010). 'The bone is mine': Affective and referential aspects of dog growls. *Animal Behaviour*, 79, 917–925. <https://doi.org/10.1016/j.anbehav.2010.01.005>
150. Sadhukhan, S., Hennelly, L., & Habib, B. (2019). Characterising the harmonic vocal repertoire of the Indian wolf (*Canis lupus pallipes*). *PLoS ONE*, 14, e0216186. <https://doi.org/10.1371/journal.pone.0216186>
151. Policht, R., Kowalczyk, A., Łukaszewicz, E., & Hart, V. (2020). Hissing of geese: Caller identity encoded in a non-vocal acoustic signal. *PeerJ*, 8, e10197. <https://doi.org/10.7717/peerj.10197>
152. Dutour, M., Lévy, L., Lengagne, T., Holveck, M.-J., Crochet, P.-A., Perret, P., Doutrelant, C., & Grégoire, A. (2020). Hissing like a snake: Bird hisses are similar to snake hisses and prompt similar anxiety behavior in a mammalian model. *Behavioral Ecology and Sociobiology*, 74, 1–9. <https://doi.org/10.1007/s00265-019-2778-5>
153. Young, B. A., Nejman, N., Meltzer, K., & Marvin, J. (1999). The mechanics of sound production in the puff adder *Bitis arietans* (Serpentes: Viperidae) and the information content of the snake hiss. *Journal of Experimental Biology*, 202, 2281–2289. <https://doi.org/10.1242/jeb.202.17.2281>
154. Kirchner, W. H., & Röschard, J. (1999). Hissing in bumblebees: An interspecific defence signal. *Insectes Sociaux*, 46, 239–243. <https://doi.org/10.1007/s000400050140>
155. Sen, S. M., Fuchs, S., Werber, C., & Tautz, R. (2002). Worker piping triggers hissing for coordinated colony defence in the dwarf honeybee *Apis florea*. *Zoology*, 105, 215–223. <https://doi.org/10.1078/0944-2006-00064>
156. Wiley, R. H., & Richards, D. G. (1978). Physical constraints on acoustic communication in the atmosphere: Implications for the evolution of animal vocalizations. *Behavioral Ecology and Sociobiology*, 3, 69–94. <https://doi.org/10.1007/BF00300047>
157. Marten, K., Quine, D., & Marler, P. (1977). Sound transmission and its significance for animal vocalization. *Behavioral Ecology and Sociobiology*, 2, 291–302. <https://doi.org/10.1007/BF00299741>
158. Marten, K., & Marler, P. (1977). Sound transmission and its significance for animal vocalization: I. Temperate habitats. *Behavioral Ecology and Sociobiology*, 2, 271–290.
159. Horowitz, S. S. (2012). *The universal sense: How hearing shapes the mind*. Bloomsbury.
160. Rothstein, S. I., & Fleischer, R. C. (1987). Vocal dialects and their possible relation to honest status signalling in the brown-headed cowbird. *The Condor: Ornithological Applications*, 89, 1. <https://doi.org/10.2307/1368756>
161. Brown, T. J., & Handford, P. (1996). Acoustic signal amplitude patterns: A computer simulation investigation of the acoustic adaptation hypothesis. *The Condor: Ornithological Applications*, 98, 608–623. <https://doi.org/10.2307/1369573>
162. Freitas, B., D'amelio, P. B., Milá, B., Thébaud, C., & Janicke, T. (2025). Meta-analysis of the acoustic adaptation hypothesis reveals no support for the effect of vegetation structure on acoustic signalling across terrestrial vertebrates. *Biological Reviews*, 100, 815–833. <https://doi.org/10.1111/brv.13163>
163. Waser, P. M., & Waser, M. S. (1977). Experimental studies of primate vocalization: Specializations for long-distance propagation. *Zeitschrift für Tierpsychologie*, 43, 239–263. <https://doi.org/10.1111/j.1439-0310.1977.tb00073.x>
164. De Kort, S. R., Elderirmire, E. R. B., Cramer, E. R. A., & Vehrencamp, S. L. (2009). The deterrent effect of bird song in territory defense. *Behavioral Ecology*, 20, 200–206. <https://doi.org/10.1093/beheco/arn135>
165. Naguib, M., Titulaer, M., Waas, J. R., Van Oers, K., Sprau, P., & Snijders, L. (2022). Prior territorial responses and home range size predict territory defense in radio-tagged great tits. *Behavioral Ecology and Sociobiology*, 76, 35. <https://doi.org/10.1007/s00265-022-03143-3>
166. Adams, E. S. (2001). Approaches to the study of territory size and shape. *Annual Review of Ecology and Systematics*, 32, 277–303.
167. Potts, J. R., & Lewis, M. A. (2014). How do animal territories form and change? Lessons from 20 years of mechanistic modelling. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20140231. <https://doi.org/10.1098/rspb.2014.0231>
168. Giuggioli, L., Potts, J. R., & Harris, S. (2011). Animal interactions and the emergence of territoriality. *PLoS Computational Biology*, 7, e1002008.
169. Stamps, J. (1994). Territorial behavior: Testing the assumptions. *Advances in the Study of Behavior*, 23, 173–232. [https://doi.org/10.1016/S0065-3454\(08\)60354-X](https://doi.org/10.1016/S0065-3454(08)60354-X)
170. Fedy, B. C., & Stutchbury, B. J. M. (2005). Territory defence in tropical birds: Are females as aggressive as males? *Behavioral Ecology and Sociobiology*, 58, 414–422.
171. Hinde, A. (1956). The biological significance of the territories of birds. *Ibis*, 98, 340–369. <https://doi.org/10.1111/j.1474-919X.1956.tb01419.x>
172. Clutton-Brock, T. H., Albon, S. D., Gibson, R. M., & Guinness, F. E. (1979). The logical stag: Adaptive aspects of fighting in red deer (*Cervus elaphus* L.). *Animal Behaviour*, 27, 211–225. [https://doi.org/10.1016/0003-3472\(79\)90141-6](https://doi.org/10.1016/0003-3472(79)90141-6)

173. Harrington, F. H., & Mech, L. D. (1979). Wolf howling and its role in territory maintenance. *Behaviour*, 68, 207–249. <https://doi.org/10.1163/156853979x00322>
174. Marshall, J. T., & Marshall, E. R. (1976). Gibbons and their territorial songs. *Science* (1979), 193, 235–237. <https://doi.org/10.1126/science.193.4249.235>
175. Smotherman, M., Knörnschild, M., Smarsh, G., & Bohn, K. (2016). The origins and diversity of bat songs. *Journal of Comparative Physiology A*, 202, 535–554. <https://doi.org/10.1007/s00359-016-1105-0>
176. Qin, H., Feng, L., Zhao, X., Sun, C., Feng, J., & Jiang, T. (2020). Great Himalayan leaf-nosed bats produce different territorial calls to respond to sympatric species and non-living objects. *Animals*, 10, 2040. <https://doi.org/10.3390/ani10112040>
177. Sun, C., Zhang, C., Lucas, J. R., Lin, A., Feng, J., & Jiang, T. (2021). Territorial calls of the bat *Hipposideros armiger* may encode multiple types of information: Body mass, dominance rank and individual identity. *Animal Cognition*, 24, 689–702. <https://doi.org/10.1007/s10071-020-01455-3>
178. Kořanová, D., Němcová, L., Policht, R., Hart, V., Begall, S., & Burda, H. (2021). Wolf howling and emergency sirens: A hypothesis of natural and technical convergence of aposematic signals. *Acta Biotheoretica*, 69, 53–65. <https://doi.org/10.1007/s10441-020-09389-6>
179. Reschke, J.-O., Krüger, S., & Hertler, C. (2024). Early hominins: Successful hunters, catchers, or scavengers? An agent-based model about hunting strategies in tropical grasslands. *Quaternary Environments and Humans*, 2, 100019. <https://doi.org/10.1016/j.qeh.2024.100019>
180. Zhao, L., Roy, S., & Wang, X. (2019). Rapid modulations of the vocal structure in marmoset monkeys. *Hearing Research*, 384, 107811. <https://doi.org/10.1016/J.HEARES.2019.107811>
181. Oliveira, D. A. G., & Ades, C. (2004). Long-distance calls in Neotropical primates. *Anais Da Academia Brasileira De Ciencias*, 76, 393–398. <https://doi.org/10.1590/S0001-37652004000200031>
182. Snowdon, C. T. (1989). Vocal communication in New World monkeys. *Journal of Human Evolution*, 18, 611–633. [https://doi.org/10.1016/0047-2484\(89\)90097-3](https://doi.org/10.1016/0047-2484(89)90097-3)
183. Bradley, C. E., & McClung, M. R. (2015). Vocal divergence and discrimination of long calls in tamarins: A comparison of allopatric populations of *Saguinus fuscicollis nigrifrons* and *S. f. lagonotus*. *American Journal of Primatology*, 77, 679–687. <https://doi.org/10.1002/AJP.22390>
184. Adret, P., Dingess, K., Caselli, C., Vermeer, J., Martínez, J., Luna Amancio, J., Van Kuijk, S., Hernani Lineros, L., Wallace, R., Fernandez-Duque, E., & Di Fiore, A. (2018). Duetting patterns of titi monkeys (Primates, Pitheciidae: Callicebinae) and relationships with phylogeny. *Animals*, 8, Article 178. <https://doi.org/10.3390/ani8100178>
185. Dolotovskaya, S., & Heymann, E. W. (2022). Coordinated singing in coppery titi monkeys (*Plecturocebus cupreus*): Resource or mate defense? *Frontiers in Ecology and Evolution*, 10, Article 898509.
186. Tilson, R. L., & Tenaza, R. R. (1976). Monogamy and duetting in an Old World monkey. *Nature*, 263, 320–321. <https://doi.org/10.1038/263320a0>
187. Gamba, M., Torti, V., Estienne, V., Randrianarison, R. M., Valente, D., Rovara, P., Bonadonna, G., Friard, O., & Giacoma, C. (2016). The Indris have got rhythm! Timing and pitch variation of a primate song examined between sexes and age classes. *Frontiers in Neuroscience*, 10, Article 249.
188. Bohr, Y. E.-M. B., Purbatrapila, A., Perwitasari-Farajallah, D., Ganzhorn, J. U., & Merker, S. (2023). Strange tunes—Acoustic variation and character displacement in a tarsier hybrid zone. *International Journal of Primatology*, 44, 581–612. <https://doi.org/10.1007/s10764-023-00351-0>
189. Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS ONE*, 1, e73.
190. Merker, B. (2000). Synchronous chorusing and the origins of music. *Musicae Scientiae*, 3, 59–73. <https://doi.org/10.1177/102986490000305105>
191. Ma, H., Wang, Z., Han, P., Fan, P., Chapman, C. A., Garber, P. A., & Fan, P. (2024). Small apes adjust rhythms to facilitate song coordination. *Current Biology*, 34, 935–945.e3. <https://doi.org/10.1016/j.cub.2023.12.071>
192. Pomberger, T., Risueno-Segovia, C., Gultekin, Y. B., Dohmen, D., & Hage, S. R. (2019). Cognitive control of complex motor behavior in marmoset monkeys. *Nature Communications*, 10, 1–8. <https://doi.org/10.1038/s41467-019-11714-8>
193. Ghazanfar, A. A., Liao, D. A., & Takahashi, D. Y. (2019). Volition and learning in primate vocal behaviour. *Animal Behaviour*, 151, 239–247. <https://doi.org/10.1016/j.anbehav.2019.01.021>
194. Liao, D. A., Zhang, Y. S., Cai, L. X., & Ghazanfar, A. A. (2018). Internal states and extrinsic factors both determine monkey vocal production. *PNAS*, 115, 3978–3983. https://doi.org/10.1073/PNAS.1722426115/SUPPL_FILE/PNAS.201722426SI.PDF
195. Koda, H., Oyakawa, C., Kato, A., & Masataka, N. (2007). Experimental evidence for the volitional control of vocal production in an immature gibbon. *Behaviour*, 144, 681–692. <https://doi.org/10.1163/15685390781347817>
196. Cowlshaw, G. (1992). Song function in gibbons. *Behaviour*, 121, 131–153. <https://doi.org/10.1163/156853992x00471>
197. Ham, S., Lappan, S., Hedwig, D., & Choe, J. C. (2017). Female songs of the nonduetting Javan gibbons (*Hylobates moloch*) function for territorial defense. *International Journal of Primatology*, 38, 533–552. <https://doi.org/10.1007/s10764-017-9964-y>
198. Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A. D., Wittig, R. M., & Crockford, C. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology*, 5, 410. <https://doi.org/10.1038/s42003-022-03350-8>
199. Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice*, 21, 531–540. <https://doi.org/10.1016/j.jvoice.2006.03.002>
200. Filippi, P. (2016). Emotional and interactional prosody across animal communication systems: A comparative approach to the emergence of language. *Frontiers in Psychology*, 07, Article 1393. <https://doi.org/10.3389/fpsyg.2016.01393>
201. Debracque, C., Slocombe, K. E., Clay, Z., Grandjean, D., & Gruber, T. (2023). Humans recognize affective cues in primate vocalizations: Acoustic and phylogenetic perspectives. *Scientific Reports*, 13, 10900. <https://doi.org/10.1038/s41598-023-37558-3>
202. Brown, S. (2000). The “musilanguage” model of musical evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). The MIT Press.
203. Brown, S. (2017). A joint prosodic origin of language and music. *Frontiers in Psychology*, 8, 1894. <https://doi.org/10.3389/fpsyg.2017.01894>
204. Fitch, W. T. (2013). Musical protolanguage: Darwin’s theory of language evolution revisited. In J. J. Bolhuis, & M. Everaert (Eds.), *Birdsong, speech, and language: Exploring the evolution of mind and brain* (pp. 489–503). The MIT Press.
205. Filippi, P. (2020). Emotional voice intonation: A communication code at the origins of speech processing and word-meaning associations? *Journal of Nonverbal Behavior*, 44, 395–417. <https://doi.org/10.1007/s10919-020-00337-z>
206. Panksepp, J. (2009). The emotional antecedents to the evolution of music and language. *Musicae Scientiae*, 13, 229–259. <https://doi.org/10.1177/1029864909013002111>
207. Slocombe, K. E., Lahiff, N. J., Wilke, C., & Townsend, S. W. (2022). Chimpanzee vocal communication: What we know from the wild. *Current Opinion in Behavioral Sciences*, 46, 101171. <https://doi.org/10.1016/j.cobeha.2022.101171>

208. Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, 61, 1203–1216. <https://doi.org/10.1006/anbe.2000.1706>
209. Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013). Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour*, 86, 189–196. <https://doi.org/10.1016/j.anbehav.2013.05.010>
210. Mitani, J. C., & Nishida, T. (1993). Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour*, 45, 735–746. <https://doi.org/10.1006/anbe.1993.1088>
211. Clark, A. P., & Wrangham, R. W. (1994). Chimpanzee arrival pant-hoots: Do they signify food or status? *International Journal of Primatology*, 15, 185–205. <https://doi.org/10.1007/BF02735273>
212. Fedurek, P., Donnellan, E., & Slocombe, K. E. (2014). Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behavioral Ecology and Sociobiology*, 68, 1345–1355. <https://doi.org/10.1007/s00265-014-1745-4>
213. Siervo, J., de Kort, S. R., & Hartley, I. R. (2023). Sexual selection for both diversity and repetition in birdsong. *Nature Communications*, 14, 3600. <https://doi.org/10.1038/s41467-023-39308-5>
214. Vallet, E., Beme, I., & Kreutzer, M. (1998). Two-note syllables in canary songs elicit high levels of sexual display. *Animal Behaviour*, 55, 291–297. <https://doi.org/10.1006/anbe.1997.0631>
215. Rossetto, F., & Laiolo, P. (2024). Potential social facilitation through song in bird communities. *Behavioral Ecology and Sociobiology*, 78, 10. <https://doi.org/10.1007/s00265-023-03427-2>
216. Loning, H., Griffith, S. C., & Naguib, M. (2024). The ecology of zebra finch song and its implications for vocal communication in multi-level societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 379, 20230191. <https://doi.org/10.1098/rstb.2023.0191>
217. Mathevon, N., Koralek, A., Weldele, M., Glickman, S. E., & Theunissen, F. E. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. *BMC Ecology*, 10, 9. <https://doi.org/10.1186/1472-6785-10-9>
218. Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates: Journal of Primatology*, 28, 199–223. <https://doi.org/10.1007/BF02382570>
219. Richman, B. (1976). Some vocal distinctive features used by gelada monkeys. *Journal of the Acoustical Society of America*, 60, 718–724. <https://doi.org/10.1121/1.381144>
220. Dobbs, O. L., Talavera, J. B., Rossi, S. M., Menjivar, S., & Gray, D. A. (2020). Signaler–receiver–eavesdropper: Risks and rewards of variation in the dominant frequency of male cricket calls. *Ecology & Evolution*, 10, 12364–12371. <https://doi.org/10.1002/ece3.6866>
221. Tuttle, M. D., & Ryan, M. J. (1981). Bat predation and the evolution of frog vocalizations in the neotropics. *Science* (1979), 214, 677–678. <https://doi.org/10.1126/science.214.4521.677>
222. Dunlop, R. A., & Noad, M. J. (2021). The eavesdropping risk of conspicuous sexual signaling in humpback whales. *Behavioral Ecology and Sociobiology*, 75, 124. <https://doi.org/10.1007/s00265-021-03048-7>
223. Osiecka, A. N., Fearey, J., Ravignani, A., & Burchardt, L. S. (2024). Isochrony in barks of Cape fur seal (*Arctocephalus pusillus pusillus*) pups and adults. *Ecology & Evolution*, 14, e11085. <https://doi.org/10.1002/ece3.11085>
224. Lord, K., Feinstein, M., & Coppinger, R. (2009). Barking and mobbing. *Behavioural Processes*, 81, 358–368. <https://doi.org/10.1016/j.beproc.2009.04.008>
225. Greene, E., & Meagher, T. (1998). Red squirrels, *Tamiasciurus hudsonicus*, produce predator-class specific alarm calls. *Animal Behaviour*, 55, 511–518. <https://doi.org/10.1006/anbe.1997.0620>
226. Leger, D. W., & Carroll, L. F. (1981). Mobbing calls of the phainopepla. *The Condor: Ornithological Applications*, 83, 377. <https://doi.org/10.2307/1367509>
227. Wheeler, B. C. (2010). Production and perception of situationally variable alarm calls in wild tufted capuchin monkeys (*Cebus apella nigrilus*). *Behavioral Ecology and Sociobiology*, 64, 989–1000. <https://doi.org/10.1007/s00265-010-0914-3>
228. Schrueth, D. M. (2024). Trophic level, arboreal position, and size of both predators and primates gave rise to a diverse and multifunctional musicality. In J. Jordania, & J. Wade (Eds.), *Defense strategies in early human evolution* (pp. 169–191). Logos.
229. Căsar, C., & Zuberbühler, K. (2012). Referential alarm calling behaviour in New World primates. *Current Zoology*, 58, 680–697. <https://doi.org/10.1093/czoolo/58.5.680>
230. Andrieu, J., Penny, S. G., Bouchet, H., Malaivijitnond, S., Reichard, U. H., & Zuberbühler, K. (2020). White-handed gibbons discriminate context-specific song compositions. *PeerJ*, 8, e9477.
231. Hattori, Y., & Tomonaga, M. (2020). Rhythmic swaying induced by sound in chimpanzees (*Pan troglodytes*). *Proceedings of the National Academy of Sciences*, 117, 936–942. <https://doi.org/10.1073/pnas.1910318116>
232. Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Scientific Reports*, 3, 1566. <https://doi.org/10.1038/srep01566>
233. De Gregorio, C., Valente, D., Raimondi, T., Torti, V., Miaretsoa, L., Friard, O., Giacomini, C., Ravignani, A., & Gamba, M. (2021). Categorical rhythms in a singing primate. *Current Biology*, 31, R1379–R1380. <https://doi.org/10.1016/j.cub.2021.09.032>
234. De Gregorio, C., Raimondi, T., Bevilacqua, V., Pertosa, C., Valente, D., Carugati, F., Bandoli, F., Favaro, L., Lefaux, B., Ravignani, A., & Gamba, M. (2024). Isochronous singing in 3 crested gibbon species (*Nomascus* spp.). *Current Zoology*, 70, 291–297. <https://doi.org/10.1093/cz/zoad029>
235. Raimondi, T., Di Panfilo, G., Pasquali, M., Zarantonello, M., Favaro, L., Savini, T., Gamba, M., & Ravignani, A. (2023). Isochrony and rhythmic interaction in ape duetting. *Proceedings of the Royal Society B: Biological Sciences*, 290, 20222244. <https://doi.org/10.1098/rspb.2022.2244>
236. Lameira, A. R., Hardus, M. E., Ravignani, A., Raimondi, T., & Gamba, M. (2024). Recursive self-embedded vocal motifs in wild orangutans. *Elife*, 12, RP88348. <https://doi.org/10.7554/eLife.88348>
237. Van Der Vleuten, B. J. R., Hovenkamp, V. A., Varkevisser, J. M., & Spierings, M. J. (2024). Context-dependent rhythmicity in chimpanzee displays. *Proceedings of the Royal Society B: Biological Sciences*, 291, 20242200. <https://doi.org/10.1098/rspb.2024.2200>
238. Muller, M. N. (2002). Agonistic relations among Kanyawara chimpanzees. In L. Marchant, C. Boesch, & G. Hohmann (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 112–124). Cambridge University Press. <https://doi.org/10.1017/CBO9780511606397.012>
239. Hosaka, K. (2015). Intimidation display. In M. Nakamura, K. Hosaka, N. Itoh, et al. (Eds.), *Mahale chimpanzees: 50 years of research* (pp. 435–447). Cambridge University Press.
240. Patel, A. D. (2021). Vocal learning as a preadaptation for the evolution of human beat perception and synchronization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376, 20200326. <https://doi.org/10.1098/rstb.2020.0326>
241. Clark, G., Saniotis, A., Bednarik, R., Lindahl, M., & Henneberg, M. (2024). Hominin musical sound production: Palaeoecological contexts and self domestication. *Anthropological Review*, 87, 17–61. <https://doi.org/10.18778/1898-6773.87.2.02>
242. Qirko, H. (2024). Pace setting as an adaptive precursor of rhythmic musicality. *Annals of the New York Academy of Sciences*, 1533, 5–15. <https://doi.org/10.1111/nyas.15120>
243. Martins, P. T., & Boeckx, C. (2020). Vocal learning: Beyond the continuum. *PLoS Biology*, 18, e3000672.
244. Tyack, P. L. (2020). A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375, 20180406. <https://doi.org/10.1098/rstb.2018.0406>

245. Janik, V. M., & Knörnschild, M. (2021). Vocal production learning in mammals revisited. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376, 20200244. <https://doi.org/10.1098/rstb.2020.0244>
246. Carouso-Peck, S., Goldstein, M. H., & Fitch, W. T. (2021). The many functions of vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376, 20200235. <https://doi.org/10.1098/rstb.2020.0235>
247. Clark, G., & Henneberg, M. (2017). *Ardipithecus ramidus* and the evolution of language and singing: An early origin for hominin vocal capability. *HOMO*, 68, 101–121. <https://doi.org/10.1016/j.jchb.2017.03.001>
248. Podlipniak, P. (2022). Pitch syntax as part of an ancient protolanguage. *Lingua*, 271, 103238. <https://doi.org/10.1016/J.LINGUA.2021.103238>
249. Becker, J. (2004). *Deep listeners: Music, emotion, and trance*. Indiana University Press.
250. McAngus Todd, N. P., & Merker, B. (2004). Siamang gibbons exceed the saccular threshold: Intensity of the song of *Hylobates syndactylus*. *Journal of the Acoustical Society of America*, 115, 3077–3080. <https://doi.org/10.1121/1.1736273>
251. Comella, I., Tasirin, J. S., Klinck, H., Johnson, L. M., & Clink, D. J. (2022). Investigating note repertoires and acoustic tradeoffs in the duet contributions of a basal haplorhine primate. *Frontiers in Ecology and Evolution*, 10, 910121.
252. Dissanayake, E. (2006). Ritual and ritualiation: Musical means of conveying and shaping emotion in humans and other animals. In S. Brown, & U. Volgsten (Eds.), *Music and manipulation: On the social uses and social control of music* (pp. 31–56). Berghahn Books. <https://doi.org/10.1515/9780857457141-007>
253. Youngerman, S. (1974). Maori dancing since the eighteenth century. *Ethnomusicology*, 18, 75. <https://doi.org/10.2307/850061>
254. Newson, M. (2017). Football, fan violence, and identity fusion. *International Review for the Sociology of Sport*, 54, 431–444. <https://doi.org/10.1177/1012690217731293>
255. Gleason, B. P. (2015). Military music in the United States: A historical examination of performance and training. *Music Educators Journal*, 101, 37–46. <https://doi.org/10.1177/0027432114563718>
256. Kartomi, M. (2024). The defense potential of music and trance according to ethnomusicological materials from the Indonesian Province of Aceh. In J. Jordania & J. Wade (Eds.), *Defense strategies in early human evolution* (pp. 84–87). Logos.
257. Wade, J. (2024). Battle trance: From ancient survival to miraculous healing and super performance. In J. Jordania & J. Wade (Eds.), *Defense strategies in early human evolution* (pp. 135–168). Logos.
258. Lehmann, C., Welker, L., & Schiefenhövel, U. W. (2008). Der Singstreit in humanethnologischer Perspektive. *Musicae Scientiae*, 12, 115–145. <https://doi.org/10.1177/102986490801200107>
259. Krause, J., & Ruxton, G. D. (2002). *Living in groups*. Oxford University Press. <https://doi.org/10.1093/oso/9780198508175.001.0001>
260. Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. The MIT Press.
261. Agus, T. R., Suied, C., & Pressnitzer, D. (2019). Timbre recognition and sound source identification. In K. Siedenburg, C. Saitis, S. McAdams, et al. (Eds.), *Timbre: Acoustics, perception, and cognition* (pp. 59–85). Springer International Publishing. https://doi.org/10.1007/978-3-030-14832-4_3
262. Oh, Y., Zuwala, J. C., Salvagno, C. M., & Tilbrook, G. A. (2022). The impact of pitch and timbre cues on auditory grouping and stream segregation. *Frontiers in Neuroscience*, 15, 725093.
263. Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, 8, 465–471. <https://doi.org/10.1016/j.tics.2004.08.008>
264. Sekulic, R. (1982). The function of howling in red howler monkeys (*Alouatta seniculus*). *Behaviour*, 81, 38–54. <https://doi.org/10.1163/156853982x00517>
265. de Cunha, R. G. T., de Oliveira, D. A. G., Holzmann, I., & Kitchen, D. M. (2015). Production of loud and quiet calls in howler monkeys. In M. M. Kowalewski, P. A. Garber, L. Cortés-Ortiz, et al. (Eds.), *Howler monkeys: Adaptive radiation, systematics, and morphology* (pp. 337–368). Springer New York. https://doi.org/10.1007/978-1-4939-1957-4_13
266. Cornick, L. A., & Markowitz, H. (2002). Diurnal vocal patterns of the black howler monkey (*Alouatta pigra*) at Lamanai, Belize. *Journal of Mammalogy*, 83, 159–166. [https://doi.org/10.1644/1545-1542\(2002\)083\(0159:DVPOTB\)2.0.CO;2](https://doi.org/10.1644/1545-1542(2002)083(0159:DVPOTB)2.0.CO;2)
267. Mitani, J. C. (1985). Responses of gibbons (*Hylobates muelleri*) to self, neighbor, and stranger song duets. *International Journal of Primatology*, 6, 193–200. <https://doi.org/10.1007/BF02693653>
268. Brown, S. (2007). Contagious heterophony: A new theory about the origins of music. *Musicae Scientiae*, 11, 3–26. <https://doi.org/10.1177/102986490701100101>
269. Nikolsky, A. (2018). Commentary: The ‘musilanguage’ model of language evolution. *Frontiers in Psychology*, 9, Article 75.
270. Lorenzi, E., Perrino, M., & Vallortigara, G. (2021). Numerosities and other magnitudes in the brains: A comparative view. *Frontiers in Psychology*, 12, 641994.
271. Bannan, N. (2012). Harmony and its role in human evolution. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 288–340). Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199227341.003.0012>
272. Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99. [https://doi.org/10.1016/S0065-3454\(08\)60377-0](https://doi.org/10.1016/S0065-3454(08)60377-0)
273. Fitch, W. T., & Jarvis, E. D. (2013). Birdsong and other animal models for human speech, song, and vocal learning. In M. A. Arbib (Ed.), *Language, music and the brain* (pp. 499–539). The MIT Press.
274. ten Cate, C. (2021). Re-evaluating vocal production learning in non-oscine birds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376, 20200249. <https://doi.org/10.1098/rstb.2020.0249>
275. Peek, F. W. (1972). An experimental study of the territorial function of vocal and visual display in the male red-winged blackbird (*Agelaius phoeniceus*). *Animal Behaviour*, 20, 112–118. [https://doi.org/10.1016/S0003-3472\(72\)80180-5](https://doi.org/10.1016/S0003-3472(72)80180-5)
276. Wheeldon, A., Szymański, P., Surmacki, A., & Osiejuk, T. S. (2021). Song type and song type matching are important for joint territorial defense in a duetting songbird. *Behavioral Ecology*, 32, 883–894. <https://doi.org/10.1093/beheco/arab030>
277. Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, 28, 48–53. <https://doi.org/10.1016/j.conb.2014.06.007>
278. Whiten, A., & van de Waal, E. (2018). The pervasive role of social learning in primate lifetime development. *Behavioral Ecology and Sociobiology*, 72, 80. <https://doi.org/10.1007/s00265-018-2489-3>
279. Gultekin, Y. B., Hildebrand, D. G. C., Hammerschmidt, K., & Hage, S. R. (2025). High plasticity in marmoset monkey vocal development from infancy to adulthood. *Science Advances*, 7, eabf2938. <https://doi.org/10.1126/sciadv.abf2938>
280. Soldati, A., Fedurek, P., Dezechache, G., Muhumuza, G., Hobaiter, C., Zuberbühler, K., & Call, J. (2025). Social and individual factors mediate chimpanzee vocal ontogeny. *Scientific Reports*, 15, 8529. <https://doi.org/10.1038/s41598-025-93207-x>
281. Merker, B. (2005). The conformal motive in birdsong, music, and language: An introduction. *Annals of the New York Academy of Sciences*, 1060, 17–28. <https://doi.org/10.1196/annals.1360.003>
282. Vernes, S. C., Kriengwatana, B. P., Beeck, V. C., Fischer, J., Tyack, P. L., Ten Cate, C., & Janik, V. M. (2021). The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society B: Bio-*

- logical Sciences*, 376, 20200236. <https://doi.org/10.1098/rstb.2020.0236>
283. Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87–90. <https://doi.org/10.1038/416087a>
284. Shannon, R. V., Zeng, F.-G. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* (1979), 270, 303–304. <https://doi.org/10.1126/science.270.5234.303>
285. Galvin, III J. J., Fu, Q.-J., & Shannon, R. V. (2009). Melodic contour identification and music perception by cochlear implant users. *Annals of the New York Academy of Sciences*, 1169, 518–533. <https://doi.org/10.1111/j.1749-6632.2009.04551.x>
286. Albouy, P., Mehr, S. A., Hoyer, R. S., Ginzburg, J., Du, Y., & Zatorre, R. J. (2024). Spectro-temporal acoustical markers differentiate speech from song across cultures. *Nature Communications*, 15, 4835. <https://doi.org/10.1038/s41467-024-49040-3>
287. Creel, S. C., & Tumlin, M. A. (2012). Online recognition of music is influenced by relative and absolute pitch information. *Cognitive Science*, 36, 224–260. <https://doi.org/10.1111/j.1551-6709.2011.01206.x>
288. Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 369–404). Academic Press. <https://doi.org/10.1016/B978-0-12-381460-9.00009-2>
289. Kluender, K. R., & Kieffe, M. (2006). Speech perception within a biologically realistic information-theoretic framework. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 153–199). Academic Press. <https://doi.org/10.1016/B978-012369374-7/50007-9>
290. Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews*, 81, 181–187. <https://doi.org/10.1016/j.neubiorev.2017.02.011>
291. Margulis, E. H. (2013). Repetition and emotive communication in music versus speech. *Frontiers in Psychology*, 4, 167. <https://doi.org/10.3389/fpsyg.2013.00167>
292. Bannan, N. (2008). Language out of music: The four dimensions of vocal learning. *The Australian Journal of Anthropology*, 19, 272–293. <https://doi.org/10.1111/j.1835-9310.2008.tb00354.x>
293. Bergman, T. J., Beehner, J. C., Painter, M. C., & Gustison, M. L. (2019). The speech-like properties of nonhuman primate vocalizations. *Animal Behaviour*, 151, 229–237. <https://doi.org/10.1016/j.anbehav.2019.02.015>
294. Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 51–87). Springer. https://doi.org/10.1007/978-1-4419-6114-3_3
295. van der Made, J., & Mateos, A. (2010). Longstanding biogeographic patterns and the dispersal of early *Homo* out of Africa and into Europe. *Quaternary International*, 223–224, 195–200. <https://doi.org/10.1016/j.quaint.2009.11.015>
296. Foister, T. I. F., Žliobaitė, I., Wilson, O. E., Fortelius, M., & Tallavaara, M. (2023). Homo heterogenus: Variability in early Pleistocene *Homo* environments. *Evolutionary Anthropology: Issues, News, and Reviews*, 32, 373–385. <https://doi.org/10.1002/evan.22005>
297. Zeller, E., Timmermann, A., Yun, K.-S., Raia, P., Stein, K., & Ruan, J. (2023). Human adaptation to diverse biomes over the past 3 million years. *Science* (1979), 380, 604–608. <https://doi.org/10.1126/science.abq1288>
298. Magill, C. R., Ashley, G. M., Domínguez-Rodrigo, M., & Freeman, K. H. (2016). Dietary options and behavior suggested by plant biomarker evidence in an early human habitat. *Proceedings of the National Academy of Sciences*, 113, 2874–2879. <https://doi.org/10.1073/pnas.1507055113>
299. Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.
300. Brown, S., Phillips, E., Husein, K., & McBride, J. (2025). Musical scales optimize pitch spacing: A global analysis of traditional vocal music. *Humanities and Social Sciences Communications*, 12, 546. <https://doi.org/10.1057/s41599-025-04881-1>
301. Chomsky, N. (1965). *Aspects of the theory of syntax*. The MIT Press. [https://doi.org/10.1016/0732-118X\(86\)90008-5](https://doi.org/10.1016/0732-118X(86)90008-5)
302. Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 36, 2346–2353. <https://doi.org/10.1121/1.1919362>
303. Wagner, B., Sturdy, C. B., Weisman, R. G., & Hoeschele, M. (2022). Pitch chroma information is processed in addition to pitch height information with more than two pitch-range categories. *Attention, Perception, & Psychophysics*, 84, 1757–1771. <https://doi.org/10.3758/s13414-022-02496-1>
304. Peter, B., Stoel-Gammon, C., & Kim, D. (2008). Octave equivalence as an aspect of stimulus-response similarity during nonword and sentence imitations in young children. In S. Maduerira, C. Reis, & P. Barbosa (Eds.), *Proceedings of the 4th international conference on speech prosody, 6–9 May, 2008*. Campinas (pp. 731–734).
305. Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. The MIT Press. <https://doi.org/10.7551/mitpress/10111.001.0001>
306. Podlipniak, P. (2025). The evolution of musicality and cross-domain co-evolutionary interactions. *Musicae Scientiae*, 29, 5–26. <https://doi.org/10.1177/10298649241255694>
307. McBride, J. M., Passmore, S., & Tlusty, T. (2023). Convergent evolution in a large cross-cultural database of musical scales. *PLoS ONE*, 18, e0284851.
308. Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & Mcdermott, J. H. (2019). Universal and non-universal features of musical pitch perception revealed by singing. *Current Biology*, 29, 3229–3243.e12. <https://doi.org/10.1016/j.cub.2019.08.020>
309. Pressnitzer, D., & Demany, L. (2019). Auditory perception: Relative universals for musical pitch. *Current Biology*, 29, R927–R929. <https://doi.org/10.1016/j.cub.2019.08.031>
310. Trehub, S. E. (2015). Cross-cultural convergence of musical features. *Proceedings of the National Academy of Sciences*, 112, 8809–8810. <https://doi.org/10.1073/pnas.1510724112>
311. Stumpf, C. (1911). *Die Anfänge der Musik [The beginnings of music]*. Johann Ambrosius Barth Verlag.
312. Nettl, B. (2005). *The study of ethnomusicology: Thirty-one issues and concepts*. New ed. University of Illinois Press Urbana.
313. Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science*, 1, 272–276. <https://doi.org/10.1111/j.1467-9280.1990.tb00213.x>
314. Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception and Psychophysics*, 52, 599–608. <https://doi.org/10.3758/BF03211696>
315. Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and brain development. *Child Development*, 58, 539. <https://doi.org/10.2307/1130197>
316. Sek, A., & Moore, B. C. J. (1995). Frequency discrimination as a function of frequency, measured in several ways. *Journal of the Acoustical Society of America*, 97, 2479–2486. <https://doi.org/10.1121/1.411968>
317. Larrouy-Maestri, P. (2018). “I know it when I hear it”: On listeners’ perception of mistuning. *Music & Science (London)*, 1, 2059204318784582. <https://doi.org/10.1177/2059204318784582>
318. Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants’ and adults’ perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 965–975. <https://doi.org/10.1037/0096-1523.25.4.965>

319. Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music*, 18, 60–72. <https://doi.org/10.1177/0305735690181005>
320. Pfordresher, P. Q., & Brown, S. (2017). Vocal mistuning reveals the origin of musical scales. *Journal of Cognitive Psychology*, 29, 35–52. <https://doi.org/10.1080/20445911.2015.1132024>
321. Pfordresher, P. Q., & Demorest, S. M. (2021). The prevalence and correlates of accurate singing. *Journal of Research in Music Education*, 69, 5–23. <https://doi.org/10.1177/0022429420951630>
322. Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception: An Interdisciplinary Journal*, 5, 1–30. <https://doi.org/10.2307/40285384>
323. Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195148367.001.0001>
324. Krumhansl, C. L. (2004). The cognition of tonality—As we know it today. *Journal of New Music Research*, 33, 253–268. <https://doi.org/10.1080/0929821042000317831>
325. Purwins, H., Blankertz, B., & Obermayer, K. (2007). Toroidal models in tonal theory and pitch-class analysis. *Computing in Musicology (Tonal Theory for the Digital Age)*, 15, 73–98.
326. Sethares, W. A. (1998). *Tuning, timbre, spectrum, scale*. Springer. <https://doi.org/10.1007/978-1-4471-4177-8>
327. Large, E. W., & Tretakis, A. E. (2005). Tonality and nonlinear resonance. *Annals of the New York Academy of Sciences*, 1060, 53–56. <https://doi.org/10.1196/annals.1360.046>
328. Large, E. W., & Almonte, F. V. (2012). Neurodynamics, tonality, and the auditory brainstem response. *Annals of the New York Academy of Sciences*, 1252, E1–E7. <https://doi.org/10.1111/j.1749-6632.2012.06594.x>
329. Kirby, S. (2002). Learning, bottlenecks and the evolution of recursive syntax. In B. Ted (Ed.), *Linguistic evolution through language acquisition* (pp. 173–204). Cambridge University Press. <https://doi.org/10.1017/CBO9780511486524.006>
330. Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107, 885–913. <https://doi.org/10.1037/0033-295X.107.4.885>
331. Opacic, T., Stevens, C., & Tillmann, B. (2009). Unspoken knowledge: Implicit learning of structured human dance movement. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35, 1570–1577. <https://doi.org/10.1037/a0017244>
332. Halpern, A. R., Zioga, I., Shankleman, M., Lindsen, J., Pearce, M. T., & Bhattacharya, J. (2017). That note sounds wrong! Age-related effects in processing of musical expectation. *Brain and Cognition*, 113, 1–9. <https://doi.org/10.1016/j.bandc.2016.12.006>
333. Schellenberg, E. G., Poon, J., & Weiss, M. W. (2017). Memory for melody and key in childhood. *PLoS ONE*, 12, e0187115.
334. Weiss, M. W., Trehub, S. E., & Schellenberg, E. G. (2012). Something in the way she sings: Enhanced memory for vocal melodies. *Psychological Science*, 23, 1074–1078. <https://doi.org/10.1177/0956797612442552>
335. Weiss, M. W., Schellenberg, E. G., & Trehub, S. E. (2017). Generality of the memory advantage for vocal melodies. *Music Perception: An Interdisciplinary Journal*, 34, 313–318. <https://doi.org/10.1525/mp.2017.34.3.313>
336. Harvey, A. R. (2017). *Music, evolution, and the harmony of souls*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198786856.001.0001>
337. Merker, B. (2019). When extravagance impresses: Recasting esthetics in evolutionary terms. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (pp. 66–86). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198804123.013.4>
338. Moscovice, L. R., Sueur, C., & Aureli, F. (2020). How socio-ecological factors influence the differentiation of social relationships: An integrated conceptual framework. *Biology Letters*, 16, 20200384. <https://doi.org/10.1098/rsbl.2020.0384>
339. Sewall, K. B. (2015). Social complexity as a driver of communication and cognition. *Integrative and Comparative Biology*, 55, 384–395. <https://doi.org/10.1093/icb/icv064>
340. Dunbar, R. I. M., & Shultz, S. (2021). Social complexity and the fractal structure of group size in primate social evolution. *Biological Reviews*, 96, 1889–1906. <https://doi.org/10.1111/brv.12730>

How to cite this article: Podlipniak, P. (2025). The role of musical pitch in long-distance defensive signaling. *Ann NY Acad Sci.*, 1550, 206–226. <https://doi.org/10.1111/nyas.70000>