



Tonal Qualia and the Evolution of Music

Piotr Podlipniak

Department of Musicology,
Adam Mickiewicz University, Poznań
podlip@poczta.onet.pl

Received 30 July 2016; accepted 30 January 2017.

Abstract

The communicative phenomena of tonal music and speech observed in all human societies differ qualitatively from other human sound expressions. This difference consists mainly of the fact that both tonal music and speech are generative, i.e., they are composed of a limited number of discrete, perceptual units organized according to some tacit rules. In the case of tonal music, these units are experienced as pitch classes ordered in time. Listening to tonally organized pitch classes leads to the experience of specific and relative emotions. These emotional qualities are called tonal qualia (often described as different tinges of tension and relaxation, or instability and stability), which are elicited thanks to our predictions concerning musical pitch structure. However, the emotional component of the recognition of musical structure seems to be incomparable with the experience of speech structure or other generative phenomena. Therefore, I propose that emotional reactions to mentally discrete pitches became an important part of the cognitive mechanism designated for music perception. I suggest that the general mechanism of prediction which acts permanently during the perception of the external world was specifically incorporated into the processing of musical pitch structure. As a result, a new, cortico-subcortical loop evolved that enables the recognition of musical pitch structure. The connection and exploitation of separate, evolutionarily old mechanisms for a new adaptive function is accordant with the evolutionary operational rule called “tinkering.” I will also discuss possible adaptive functions of pitch structure and the evolution of the ability to recognize tonal qualia by means of the Baldwin effect.

Keywords: music; tonality; qualia; evolution; cortico-subcortical loops.

Introduction

The main function of hearing is to detect information about the external world based on detectable acoustic signals (Schnupp et al., 2011). Animals are able to infer various pieces of information from acoustic waves, such as the size and shape of the object generating the sound, or the speed of an object in the environment. In addition, over the course of evolution sound has become an important tool of communication among individuals (Horowitz, 2012). As different species of animals have their own evolutionary history, their individual auditory systems, including their sound-producing organs, are characterized by specific features. As a result, many species-specific forms of sound communication now exist in the biosphere. Humans also use hearing to interpret the external world, both animate and inanimate, according to human-specific rules. As social animals, humans depend especially on the behavior of others (Wilson, 2012). They must therefore exchange information between each other, and so sound seems to be a very important medium for our communication. From a human point of view, the perception of sounds leads to a spectrum of sensations that becomes part of our internal and conscious world, thus allowing us to navigate better in our physical and social environment. One such sensation is pitch. What is especially interesting is that our perception of pitch in tonal music is characterized by a set of specific features which are absent from our perception of other sound stimuli. This specificity triggers the question about the possible biological function of musical pitch and its evolutionary sources.

The Human Sensation of Pitch

In order to understand the characteristics of tonal qualia, we must first consider the specificity of the human sensation of pitch. In terms of our auditory perception, everything that we perceive does not exist in the external world. However, there are some observable (sometimes even measurable) correlations between some objective acoustical parameters and the characteristics of these auditory categories. The crucial acoustical parameter which influences the perception of pitch is the frequency of sound waves (Roederer, 2009), although our sensation of pitch also depends to some extent on the pressure and the waveform of the sound (Stainsby & Cross, 2009). Pitch as a psychoacoustic percept is unidimensional and is defined as “. . . that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high” (Stainsby & Cross, 2009, p. 47; see also ANSI, 1994). Although pitch is conceptualized by means of different metaphors by different cultures (Zbikowski, 2002), the sensation of pitch as a conscious experience seems to be either a prelinguistic default label of various sources of sound (Huron, 2016, p. 39), or a preconceptual tool of communication (or both). This kind of sensation is not only an important part of the perception of music; it also accompanies our experience of all complex periodic sounds. These sounds are produced by the human larynx while we are singing and also during crying, laughing and speaking. Therefore, pitch as a unidimensional, psychoacoustic percept is not solely a musical domain.

Whilst listening to music, the frequency of sound waves is interpreted by the human nervous system in a complex and elaborate way. As a result of this processing of music, our experience of sound is enriched by an additional component that is often called the circular or chromatic dimension of pitch (Deutsch et al., 2008; Deutsch, 2010). The main characteristic of this dimension is the phenomenon of octave equivalence, which refers to the perceptual equivalence of sounds separated by one or more octave, i.e., the musical interval in which the fundamental frequency of the sound octave above is twice the frequency of the original sound. Based on this perceptual equivalence, the spectrum of frequencies between each octave is divided into a culture-specific set of pitch categories called “pitch class,” or “pitch chroma.” Similar to phonemes, the division into pitch classes is cognitive in nature, meaning that sounds with frequencies belonging to a zone of frequencies are recognized as the same pitch class (Rakowski, 2009). What is important is that at this level pitch is interpreted as a discrete unit (Merker, 2002). It is for precisely this reason that Zbikowski (p. 71) suggests that because the division of the frequency spectrum into discontinuous units is similar to the division of the spatial domain into points, the metaphors used to describe pitch are often spatial. Cox (2011), by contrast, claims that these metaphors are grounded in our bodily experience of music. Whatever the actual reason for these metaphors, the necessity of describing pitch in a metaphorical way suggests by itself that pitch is a preconceptual phenomenon. Moreover, the interpretation of periodic sounds in terms of pitch classes is not obligatory and depends on some additional conditions. This cognitive nature of pitch class recognition is well illustrated by a “speech to song” illusion in which the same acoustical stimulus is previously interpreted as speech and then as a sequence of musical pitch classes (Deutsch et al., 2011). Although octave equivalence has been observed among certain mammalian species of dolphins (Richards et al., 1984) and monkeys (Wright et al., 2000), the use of pitch classes as the basis for a syntactically complex means of communication (tonal music) seems to be a human-specific domain.

Tonal Qualia as a Specific Interpretation of Pitch

The experience of pitch in tonal music is characterized by yet another, additional cognitive dimension which seems quite different to the two previously discussed. This section is devoted to the contemporary knowledge regarding the mechanisms which are responsible for this additional interpretation, as well as the specific characteristics of this dimension. Every pitch class experienced as part of a tonal melody is perceived as possessing some additional sensual quality depending on the tonal function of a particular pitch class, which is fulfilled in the structure of the melody. For example, the quality of the most important tonal function known as “tonic” is often described as “stable,” “resolved,” “contented” etc. (Huron, 2006, p. 145). On the other hand, pitch classes fulfilling quite different functions such as the so-called “leading tone” are described as “unstable,” “restless” etc. (p. 145). These sensual qualities are usually referred to as “tonal qualia” (Huron, 2006; Margulis, 2013) and are intersubjectively recognizable by people enculturated in the same musical culture. While the sensation of pitch as a psychoacoustic percept and an isolated

pitch class is relatively invariant, i.e., the same physical properties of sound (if interpreted as music) cause similar qualia, the experience of tonal qualia depends on the context of other pitch classes in which a particular pitch class is discerned (Huron, 2006). In other words, the same physical properties of sound may cause different sensations depending solely on the musical context of a perceived sound (note). Although in both Western and non-Western artistic music the sensations of tonal qualia can be influenced by many additional factors (e.g., harmony in the case of Western music), from an evolutionary point of view tonality should be considered in the broader sense as a system of horizontal relations between pitch classes, rather than complex tonal harmony, which is its Western cultural elaboration (Jackendoff & Lerdahl, 2006), consisting of the statistical patterning of chords (McMullen Jonaitis & Saffran, 2009) instead of single pitch classes.

There is also another difference between qualia experienced as a result of the interpretation of pitch as a psychoacoustic percept and as a tonal function. All such tonal qualia are described in terms of particular emotional qualities. According to some scholars, one should differentiate between “basic qualia” such as size, texture, color and consonance, and “emotionally loaded qualia,” which are experienced as if they possess an intrinsic emotional tinge (Huston & Huston, 2015, p. 484). What is equally interesting and puzzling is the fact that there is nothing resembling tonal qualia in terms of their emotional qualities in other domains of human cognition. Although there are some similarities, e.g., between the perception of tonal center (in the auditory domain) and focal colors or horizontal, vertical, and 45° angled lines (Rosch, 1975) (in the visual domain), which are both treated as a cognitive reference point (Bharucha, 1984), the experience of tonal qualia seems to be unique in that it serves as the preconceptual basis of an intuitive, generative, and communicative system. This basis is a tonal hierarchy which is in fact a kind of tacit, emotionally founded set of relations between pitch classes. Although phonemes (Pierrehumbert, 2003), similar to pitch classes (Ong et al., 2017), are learned by means of distributional learning, the qualia of phonemes are experienced differently without anything resembling pitch center or pitch hierarchy. The exceptionality of tonal qualia in this respect suggests that their domain-specificity is strictly connected to human musicality. In other words, tonal qualia are unique, music-specific phenomena in the same way as the sensations of phonemes or words are unique and specific to speech.

There is yet another characteristic of tonal qualia which causes them to be an indispensable part of our experience of tonal music. The sensation of different qualities depending on pitch class context seems to be the basis for the recognition of pitch syntax. Like language, music is an example of the Humboldt system (Merker, 2002), meaning that both sentences and melodies are based on a restricted number of units composed according to certain rules. As a result, one can generate an infinite number of correct sentences and melodies but not all possible sentences and melodies are correct. In order to be correct, some pitch classes (or words and/or phonemes in the case of language) should be placed correctly. People intuitively recognize whether a particular phrase (sentence or melody) is correct or not despite usually being not aware of the particular rules being fulfilled or not in the recognized phrase (Tillmann et al., 2000). This means that we have certain expectations.

As people listening to music have expectations about which pitch class is more likely to occur, we can infer that there must be a kind of pitch class hierarchy stored in the nervous system (Krumhansl, 2004). In other words, the syntactic organization of sounds imparts hierarchical qualities to the experience of these sounds that are in fact the aforementioned tonal qualia. However, while in the case of language the recognition of correctness is possible thanks to the conceptual meaning, our recognition of syntax in music seems to be based on preconceptual phenomena. Because people in all known cultures organize pitch classes according to more or less complex syntactical rules, it is very probable that pitch syntax, similar to language syntax, is a result of the human-specific propensity to communicate using pitch. It is therefore reasonable to look for the origin of pitch syntax and tonal qualia in biology.

The Biological Roots of Tonal Qualia

It is part of our nature to experience many different sensations which must have evolved as a result of natural selection. Therefore, the understanding of tonal qualia from a biological perspective requires the knowledge of a possible biological function of these sensations, as well as answering how something as unique as tonal qualia could have evolved. According to W. Tecumseh Fitch (2015), the biology of music should be investigated on all four of Tinbergen's explanatory levels, i.e., mechanistic, ontogenetic, phylogenetic and functional (1963). The first two represent the so-called "proximate explanations," while the latter two represent the "ultimate explanations." The proximate explanations answer the question regarding how something works, whereas the ultimate explanation explains why it works.

As far as the proximate explanation of tonality is concerned, the majority of contemporary theories concentrate on the mechanistic level. The diversity of experienced tonal qualia is explained mainly by means of general cognitive mechanisms that lead to either the "exposure effect" (Krumhansl, 1990) or the "prediction effect" (Huron, 2006). The exposure effect is a well-known mechanism (Zajonc, 1968) that refers to a situation in which stimuli that occur more frequently (pitch classes in our case) are assessed more positively. However, because not all the pitch classes that occur most frequently are experienced as emotionally positive, David Huron (2006) proposed that tonal qualia are the results of our expectations and therefore depend on the probability of occurrence. In this explanation, the feelings elicited by our fulfilled or unfulfilled (to varying degrees) predictions of subsequent pitch classes are misattributed to these pitch classes. According to these two theories, our minds, due to statistical learning, create tonal hierarchies based on the statistical distribution of pitch classes in music that we hear throughout our lives. The ultimate reason for tonal hierarchy is claimed to be related to the adaptive value of mental mechanisms involved in expectation (Huron, 2006). Hence, in both cases tonal qualia are treated as the byproduct of general cognition. However, this view is in stark contrast to the uniqueness of tonal qualia in the variety of our sensations. In fact, tonal qualia are absent from speech, laughter, crying etc. and it is impossible to "translate" them into other modalities (e.g., gestures, pictures). Nonetheless, many phenomena composed of discrete units are learned

by statistical learning. A good example is the learning of our mother tongue, in which phoneme distribution is just as statistically stable as in the case of pitch classes in music. Yet there is nothing in our experience of phonemes that resembles the hierarchy based on different emotional tinges similar to tonal hierarchy.

Therefore, although both statistical learning and the mechanism of prediction are definitely important parts of our ability to recognize tonal qualia, they alone cannot explain the specificity of this ability. Moreover, the view that the processing of musical syntax is related to cortical activity seems to be insufficient; something more is needed to cause the different sensations of pitch classes in different contexts. The fact that the activation of subcortical structures (Koziol & Budding, 2009) is involved in statistical learning, the mechanism of prediction, and emotional reactions to tonal relations suggests that these structures must play an important role in the recognition of tonal relations. This notion contradicts the traditional point of view that structural features are mainly processed in the neocortex (Maess et al., 2001). The emotional experience as a preconceptual clue to indicate tonal function leads to the assumption that the recognition of tonal relations is possible thanks to coordinated activity of both cortical and subcortical areas of the brain. This assumption is convergent with the idea of cortico-subcortical loops as the structural basis of human cognition (Alexander et al., 1986; Gorzelańczyk, 2011). In fact, some research indicates that subcortical structures are also involved in the processing of musical syntax (Mikutta et al., 2015). Open questions remain regarding which particular brain structures are parts of the “tonal” loop and whether this loop is domain-specific, i.e., only involved in the processing of tonal structure.

Another level of proximal explanations is ontogenesis, in which the sense of key develops relatively late, i.e., around 5 years of age (Trainor & Trehub, 1994). However, even among infants a few days old, the processing of tonally simple musical stimuli is lateralized in the brain, in contrast to the processing of tonally ambiguous stimuli (Perani et al., 2010). This fact suggests that the ability to recognize tonal relations starts to develop very early in life, although it needs a lot of time to reach operational maturity. There is still the question of whether there are critical periods in the development of the ability to recognize tonal relations (Trainor, 2005). Nevertheless, the development of human abilities necessary to understand pitch syntax is composed of sequential stages (McMullen & Saffran, 2004; Brandt et al., 2012), which suggests some kind of developmental predisposition.

More puzzling is the ultimate level of explanations concerning tonal qualia. In the case of a potential adaptive function or functions related to the use of pitch classes, not one satisfactory explanation has been put forward. On the one hand, musical syntax (pitch syntax too) is treated as a byproduct of language syntax (Patel, 2003, 2008). Such a view does not explain why pitch—not, for example, timbre or dynamics—has become the basis of musical syntax. On the other hand, pitch classes are often indicated as music-specific features (e.g., Jackendoff & Lerdahl, 2006; Fitch, 2013) that necessitate an adaptive explanation if music is a biological adaptation. Many theories concerning the biological function of music

have been proposed since the dawn of Charles Darwin's theory of evolution. Darwin suggested that music evolved thanks to sexual selection (Darwin, 1871). In this theory, which has lately regained interest (Miller, 2000), music is described as being an effective tool to facilitate mating. The authors of more recent theories have postulated other adaptive functions of music: facilitating social consolidation (Roederer, 1984; Storr, 1993), informing individuals about social cohesion (Hagen & Bryant 2003), sustaining mother–infant bonds (Papoušek, 1996; Dissanayake, 2008), etc. However, music is a very complex phenomenon which depends on many abilities and is composed of various elements such as rhythm, pitch, dynamics, and timbre. All these elements can be used separately and in a culture-specific way. Therefore, the main question is which musical features are responsible for a particular adaptive function of music. The popular answer to this question concentrates on musical pulse, which allows movement to be synchronized with music (McNeill, 1995; Reddish, 2013). Although there are more and more studies suggesting that music can be used as a tool of social consolidation (Pearce et al., 2015; Weinstein et al., 2016), the view that the consolidating power of music is related solely to musical rhythm seems incomplete. After all, the majority of musical expressions, especially tribal singing, which most probably resembles the oldest form of music (Morley, 2013), consists not only of rhythms, but also of pitches. Therefore, there is the question of why people expend their energy and time in order to generate and control pitch during singing.

An interesting notion that can help to explain the consolidating function of tonal qualia is the idea of brain state alignment (Bharucha et al., 2012). According to this view, the brain states of co-performers (and listeners) align during a collective musical performance as a result of the processing of similar mental categories. From this perspective, the recognition of pitch classes as part of an unfolding pitch sequence leads to similar activities of cortical and subcortical brain regions. Because the processing of sequential stimuli seems to always depend not only on cognitive processes, but also on a form of motor control (Schulkin, 2013, p. 75), it is reasonable to assume that the recognition of pitch syntax is partly motor in nature too. Additionally, the intuitive recognition of pitch structure is facilitated by emotional prompts (feelings of uncertainty – structural incompleteness; feelings of relaxation – structural completeness). Therefore, listening to intersubjectively recognizable tonal qualia by a group of people manifests in the experiencing of similar motions, emotions and perceptual categories. Such an alignment of motor, emotional and cognitive brain states causes feelings of group identity and facilitates mutual trust (Bharucha et al., 2012; Podlipniak, 2016). From this point of view, tonal qualia are intersubjective properties of not only the spectral and emotional, but also the motional synchronization of human brains. As such, they may be described from the perspective of embodied music cognition. According to the mimetic hypothesis (Cox, 2011, 2016), musical imagery is partly motor in nature and consists of “physical empathy.” This means that vocal communication at least partly involves “mimetic cognition,” which is strictly related to imitation, not only as a method of learning, but also as a means of comprehending meaning. This view is especially relevant to the theories which emphasize the role of vocal learning and ritual culture in the evolution of language and music (Merker, 2005).

Because vocal learning is in fact a kind of learning by imitation that necessitates the motor control of the larynx, pitch class comprehension must be somehow related to motor schemata which mentally represent the production of pitches by the vocal apparatus. Therefore, the aforementioned alignment of brain states (Bharucha et al., 2012) can be understood in terms of “mimetic engagement” (Cox, 2011). However, the framework used in embodied music cognition needs a clear hypothesis as to how exactly the body influences the processing of music (Matyja, 2016).

The final problem concerning the biological explanation of tonal qualia is the phylogenesis of the ability that enables the experience of tonal relations. In other words, how and when did the ability to recognize tonal qualia evolve? Chimpanzees—our closest animal relatives—do not use pitch classes as a way of communication, which suggests that the ability to recognize tonal qualia must have evolved during our evolutionary lineage after the split from our common ancestor with the chimpanzee. The species-specific character of pitch syntax as a part of vocal expression indicates that tonal qualia probably played an important role in our cognition no earlier than the vocal control of pitch. The evolution of vocal control was the necessary change to allow vocal learning, i.e., the ability of the vocal apparatus to reproduce what is heard (Merker, 2012). It is assumed that the control of the larynx which allows the ability to sustain the generation of a particular frequency of sound is no older than 600,000 years (Morley, 2013). What is extremely interesting is that this ability is only necessary to sing, but not to speak (Bannan, 2012). Of course, the ability to sustain or control pitch is not the same as the ability to organize pitches syntactically. After all, many animal species are able to sustain and control sound frequency with their vocal organs, but they do not communicate by means of pitch syntax. The appearance of pitch syntax as a part of our species-specific behavior had to be a result of a change in our brain connections. What was needed was the specific connection between emotional and motor processing, an unconscious statistical analysis of pitch class occurrence, and working memory. It seems that these elements play an exceptional role in the processing of pitch structure by humans. First of all, during singing the motor control of the larynx which allows precise tuning and helps to sustain F_0 (Bannan, 2012) necessitates accurate predictions concerning the next pitches to sing. In order to achieve this goal effectively, both statistical analysis of pitch class occurrence and the privileged place for pitch classes in working memory seem to be very useful. Additionally, emotions as a motivational mechanism are the best tool to induce human singing and reinforce learning of a particular musical idiom, especially at the pre-conceptual stage of human evolution. The evolution of a new functional connection between previously existing parts is consistent with the evolutionary operational rule called “tinkering” (Jacob, 1977). Such a functional connection would have been achieved by the evolution of new cortico-subcortical loops. The possible evolution of new neural connections functionally involved in pitch class processing could be related to vocal learning and the control of the larynx (Fitch & Jarvis, 2013). One of the differences between vocal-learning and non-vocal-learning primates is the existence of more elaborate cortico-subcortical pathways in humans in comparison to other primates (Jarvis, 2007), thus indicating the possible functional role of these connections in the processing of speech and singing,

which are human-specific vocalizations. Taking into account the fact that the proposed synchronization between human brains during consolidative rituals consists of not only spectral but also emotional and motor representations, it is probable that different cortico-striatal-thalamic-cortical circuits (especially limbic and dorsolateral prefrontal circuits) are involved in varying degrees in this process.

Another problem is the scenario of the evolution of these specific cortico-subcortical loops. Pitch syntax as a tool of social consolidation works only if a group of people use sequences of pitches according to some rules. However, evolutionary change starts as an accidental event, e.g., the mutation or recombination of DNA that occurs in one individual. Therefore, how do we explain the adaptive value of the ability to recognize tonal qualia in the case of an individual who is able to recognize that something is absent from his or her environment? In such circumstances, this ability would never proliferate in the whole population. A possible solution is the Baldwinian mode of evolution (Baldwin, 1896), which starts with a behavioral trait being invented due to cultural flexibility. This trait of phenotypic adaptation must then last many generations and be useful and biologically costly. According to Baldwin, if the aforesaid conditions are fulfilled, sooner or later this trait will start to be genetically controlled. Such a scenario has been suggested as the origin of pitch centricity (Podlipniak, 2016), which consists of preferential use of one pitch class in a sequence of pitches. The preferred pitch class—pitch center—is an example of a tonal quale. If this scenario is correct, then it is possible that other tonal qualia also became part of human-specific means of communication. From the perspective of our experiences, tonal qualia inform us that a heard pitch sequence is correct or not. From the point of view of ultimate explanations, they serve as clues to indicate social acceptance and level of consolidation within a group.

Conclusion

The proposed origin of tonal qualia as a music-specific phenomenon does not explain all issues related to the evolution of music. This view is a concept behind the evolution of a particular musical trait rather than music as a whole phenomenon. However, in contrast to the majority of music origin theories, the presented idea focuses on the possible functional specificity of pitch syntax that either has been neglected or treated as a byproduct of non-music-specific abilities. Although pitch syntax is usually only a part of a more complex syntactic structure of music, it does reveal some kind of independency. After all, there are examples of music without a measurable rhythm that still possess tonal relations. Therefore, it is reasonable to understand the ability to recognize tonal qualia as a separate, domain-specific mental tool which has evolved due to its consolidating function.

Acknowledgments

I would like to thank the reviewers for their critical and inspiring suggestions. I would also like to thank Peter Kośmider-Jones for his language consultations and patience.

References

- ANSI (1994). *ANSI S1.1-1994: American national standard on acoustical terminology*. New York, NY: American National Standards Institute.
- Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual review of neuroscience*, *9*, 357–381. doi:10.1146/annurev.ne.09.030186.002041
- Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist*, *30*(354), 441–451.
- Bannan, N. (2012). Harmony and its role in human evolution. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 288–339). Oxford, UK: Oxford University Press.
- Bharucha, J. J., Curtis, M. E., & Paroo, K. (2012). musical communication as alignment of brain states. In P. Rebuschat, M. Rohrmeier, & J. A. Hawkins (Eds.), *Language and music as cognitive systems. [volume based on an eponymous conference, Cambridge, 11–13 may 2007]* (pp. 139–155). New York, NY: Oxford University Press.
- Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in psychology*, *3*, 327. doi:10.3389/fpsyg.2012.00327
- Darwin, C. R. (1871). *The descent of man and selection in relation to sex*. London, UK: John Murray.
- Deutsch, D. (2010). The paradox of pitch circularity. *Acoustics Today*. (7), 8–14.
- Deutsch, D., Dooley, K., & Henthorn, T. (2008). Pitch circularity from tones comprising full harmonic series. *The Journal of the Acoustical Society of America*, *124*(1), 589–597. doi:10.1121/1.2931957
- Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *The Journal of the Acoustical Society of America*, *129*(4), 2245–2252. doi:10.1121/1.3562174
- Dissanayake, E. (2008). If music is the food of love, what about survival and reproductive success? *Musicae Scientiae*, *12*(1 Suppl), 169–195. doi:10.1177/1029864908012001081
- Dor, D. (2000). From the autonomy of syntax to the autonomy of linguistic semantics: Notes on the correspondence between the transparency problem and the relationship problem. *Pragmatics & Cognition*, *8*(2), 325–356.
- 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). Cambridge, MA: MIT Press.
- Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *370*(1664), 20140091. doi:10.1098/rstb.2014.0091
- Gorzelańczyk, E. J. (2011). Functional anatomy, physiology and clinical aspects of basal ganglia. In J. F. P. Peres (Ed.), *Neuroimaging for clinicians: Combining research and practice* (pp. 89–106). London, UK: InTech. doi:10.5772/24847
- Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*(1), 21–51. doi:10.1007/s12110-003-1015-z
- Horowitz, S. S. (2012). *The universal sense: How hearing shapes the mind*. New York, NY: Bloomsbury.
- Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

- Huston, A. J., & Huston, J. P. (2015). Aesthetic evaluation of art: A formal approach. In J. P. Huston, M. Nadal, F. Mora, L. F. Agnati, & C. J. C. Conde (Eds.), *Art, Aesthetics, and the Brain* (pp. 479–499). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199670000.003.0024
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York, NY: Oxford University Press.
- Krumhansl, C. L. (2004). The cognition of tonality – as we know it today. *Journal of New Music Research*, 33(3), 253–268. doi:10.1080/0929821042000317831
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4(5), 540–545. doi:10.1038/87502
- Margulis, E. H. (2013). *On repeat: How music plays the mind*. New York, NY: Oxford University Press.
- McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception*, 21(3), 289–311. doi:10.1525/mp.2004.21.3.289
- McNeill, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Cambridge, MA: Harvard University Press.
- Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae*, 6(1), 3–21. doi:10.1177/102986490200600101
- Merker, B. (2012). The vocal learning constellation: Imitation, ritual culture, encephalization. In N. Bannan (Ed.), *Music, language, and human evolution* (pp. 215–260). Oxford, UK: Oxford University Press.
- Mikutta, C. A., Dürschmid, S., Bean, N., Lehne, M., Lubell, J., Altorfer, A., Parvisi, J., Strik, W., Knight, R. T., & Koelsch, S. (2015). Amygdala and orbitofrontal engagement in breach and resolution of expectancy: A case study. *Psychomusicology: Music, Mind, and Brain*, 25(4), 357–365. doi:10.1037/pmu0000121
- 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). Cambridge, MA: MIT Press.
- Morley, I. (2013). *The prehistory of music: Human evolution, archaeology, and the origins of musicality*. Oxford, UK: Oxford University Press.
- Papoušek, H. (1996). Musicality in infancy research: Biological and cultural origins of early musicality. In I. Deliège & J. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence* (pp. 37–55). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780198523321.003.0002
- Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science*, 2(10), 150221. doi:10.1098/rsos.150221
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10), 4758–4763. doi:10.1073/pnas.0909074107

- Podlipniak, P. (2016). The evolutionary origin of pitch centre recognition. *Psychology of Music, 44*(3), 527–543. doi:10.1177/0305735615577249
- Rakowski, A. (2009). The domain of pitch in music. *Archives of Acoustics, 34*(4), 429–443.
- Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: Synchrony, shared intentionality and cooperation. *PloS one, 8*(8), e71182. doi:10.1371/journal.pone.0071182
- 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*(1), 10–28. doi:10.1037/0735-7036.98.1.10
- Roederer, J. G. (1984). The search for a survival value of music. *Music Perception: An Interdisciplinary Journal, 1*(3), 350–356. doi:10.2307/40285265
- Roederer, J. G. (2009). *The physics and psychophysics of music*. New York, NY: Springer US.
- Schnupp, J., Nelken, I., & King, A. (2011). *Auditory neuroscience: Making sense of sound*. Cambridge, MA: MIT Press.
- Schulkin, J. (2013). *Reflections on the musical mind: An evolutionary perspective*. Princeton, NJ: Princeton University Press.
- Stainsby, T., & Cross, I. (2009). The perception of pitch. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford library of psychology: The Oxford handbook of music psychology* (pp. 47–58). Oxford, NY: Oxford University Press.
- Storr, A. (1993). *Music and the mind*. New York, NY: Ballantine.
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*(4), 885–913. doi:10.1037//0033-295X.107.4.885
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie, 20*(4), 410–433. doi:10.1111/j.1439-0310.1963.tb01161.x
- Trainor, L. J. (2005). Are there critical periods for musical development? *Developmental psychobiology, 46*(3), 262–278. doi:10.1002/dev.20059
- Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics, 56*(2): 125–132. doi:10.3758/BF03213891
- Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Singing and social bonding: changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior, 37*(2), 152–158. doi:10.1016/j.evolhumbehav.2015.10.002
- Wilson, E. O. (2012). *The social conquest of earth*. New York, NY: Liveright Publishing Corp.
- Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General, 129*(3), 291–307. doi:10.1037/0096-3445.129.3.291
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1–27. doi:10.1037/h0025848