

PIOTR PODLIPNIAK

Adam Mickiewicz University in Poznań, Poland
Institute of Musicology

THE NEO-PYTHAGOREAN VIEW OF MUSICAL STRUCTURE IN THE LIGHT OF MUSIC PSYCHOLOGY

The aim of this paper is to show why the neo-Pythagorean claims concerning musical structure are out-of-date and require the incorporation of contemporary psychological knowledge. The neo-Pythagorean view of musical structure has been analyzed and confronted with the contemporary neuropsychological view of music perception. It has been also suggested that musical intervals exist solely in human brains as a kind of interpretation of acoustic sounds. These sounds can be interpreted differently depending on many factors, which the popular speech-to-song illusion clearly illustrates. Another example of neo-Pythagorean ideas about musical structure that need psychological knowledge is tonal hierarchy, which also exists solely in human brains. Therefore, the popular musicological description of musical intervals in terms of mathematical proportions is misleading. It has been proposed that current musicological theories should always be confronted and consistent with contemporary psychological knowledge. This implies closer cooperation between musicology and the psychology of music.

Keywords: musical structure; neo-Pythagoreanism; tonality; mental representations of music; musicology.

INTRODUCTION

Since antiquity the Western view about music has been strongly influenced by the Pythagorean vision of musical systems (Caleon & Ramanathan, 2008). In general, this vision is based on the idea that a musical system is composed of

Adress for correspondence: PIOTR PODLIPNIAK—Adam Mickiewicz University in Poznań, Institute of Musicology, ul. Uniwersytetu Poznańskiego 7, 61–614 Poznań, Poland; e-mail: podlip@poczta.onet.pl

music intervals reflecting mathematical proportions (Helmholtz, 1954). Although the claims concerning which physical features these proportions actually refer to—whether they are the proportions of string lengths (Crocker, 1964), the proportions of the fundamental frequencies of harmonic sounds (Papadopoulos, 2002), or proportions of the frequencies of neuronal firing (Large & Almonte, 2012; Large & Tretakis, 2005)—have been changing throughout the history of Western thought about music (Nolan, 2002), the belief in precise tuning as a foundation of music is still present in contemporary musicology as well as in other scholarly traditions related to music (for detailed examples, see Parncutt & Hair, 2018). Mathematical proportions are also used as the exact image of rhythm measure durations, which is a counterpart of pitch structure in the domain of musical time. This view is often strengthened by the widespread conviction that music is solely a physical phenomenon existing independently of human cognition in the external world i.e., outside the human brains. After all, our entire knowledge about the external world comes from our senses. Since our conviction about the reality of perceived objects allows us to survive and navigate in the environment, it is not surprising that songs and other musical pieces—the outcomes of our musical activities such as singing and playing musical instruments—are often believed to be independent of our perception, i.e., regarded as an object of perception and cognition (Kivy, 1990, p. 41). Such beliefs—i.e. the conviction that every musical structure exists independently of human perception and that this structure is based on simple mathematical proportions—I propose to name “neo-Pythagorean.”

Interestingly, such neo-Pythagorean claims are not restricted to Platonically oriented scholars. Nicolas Cook, for instance, proved that large-scale tonal relations are not audible, and he has claimed that they are “imperceptible aspects of musical structure” (Cook, 1987, p. 205). Such a claim implicitly suggests that musical structure exists independently of our perception. It is worth mentioning, however, that there are also constructivists who treat musical mental patterns such as the mental images of musical phrases and melodies as parts of music. Nonetheless, even among these scholars the neo-Pythagorean influences are often present. A good example of such a scholar is Roger Scruton. On the one hand, he indicates that hearing music means hearing a sound organization “that is not reducible to any properties of the sound, nor to any properties of the vibration that causes it” (Scruton, 1999, p. 39). On the other hand, Scruton emphasizes that pitch, as a secondary quality of sound, is reducible to vibration. As a consequence, the hearing of pitches “exactly parallels the physical order which produces them” (p. 21). Moreover, he also suggests that these secondary qualities of

sound are what humans and animals hear when they hear sounds (Scruton, 1999), which means that during perception we have access to the physical properties of sounds, the mental representations of which are in fact the building blocks of musical structure.

However, the contemporary understanding of music perception and the processing of music stimuli by the nervous system suggests that none of the crucial elements of musical structure—i.e., pitch intervals (Parncutt & Hair, 2018), metrical grouping (London, 2012), tonal relations, and rhythm measures—exists as an acoustic phenomenon (Huron, 2006). Instead, all these features are in fact constructed in the process of perception according to previously acquired tacit knowledge. This construction is based on the species-specific characteristics of our physiology as well as on the culture-specific traits of musical behavior, both of which influence the interpretation of perceived sounds by our minds. Therefore, although all these features depend on the acoustic properties of perceived sounds, they cannot be reduced to or equalized with them. In fact, all these aforesaid features of musical structure are solely mental patterns. For instance, the perception of a fifth is in fact an interpretation of sound frequencies as a mental pattern of a fifth which depends on many factors (the musical context, the listener's musical experience, intonation standards, the musical system which the listener is familiar with, etc.), rather than the recognition of the simple mathematical $3/2$ proportion of sound frequencies, as it is often misleadingly explained. This observation raises the question of what the actual object of musical analysis is.

The aim of this paper is to present and advocate the view that musical structure should be understood as a mental phenomenon that is a result of sound interpretation by the nervous system. This does not mean that the mental categories of musical structure are independent of acoustic parameters. For instance, as research following the paradigm of correlative psychoacoustic has proven, it is the correlation between the overlap of the harmonics of two simultaneously sounding sounds and the experience of sensory consonance (cf. Sethares, 1998). However, although sensory consonance can play an important role in the musical structure, for example by means of strengthening the sense of completeness in a cadence, there is no such a precise stable correlation between the harmonics of two sounds and the recognition of a musical interval. Therefore, the author argues that the mental categories of musical structure such as pitch intervals cannot be reduced to acoustic parameters or to the relations between these parameters, as the neo-Pythagoreans claim they can be.

This article starts with a short description of the cognitive view of perception. After that, the main part deals with the understanding of the term “musical structure,” which is presented together with the premises that support the proposed thesis about the mental character of musical structural features. Following on from this, the question about the objective criteria of musicality in the light of the previously proposed understanding of musical structure is raised and discussed. The article is completed by a short conclusion in which the main proposed claims about musical structure are recapitulated.

PERCEPTION AS AN INTERPRETATION

From the cognitivist point of view, the process of perception as a whole is an interpretation of the external world rather than just the reflection of its physical properties (Raftopoulos, 2009). For instance, in terms of the processing rules, the recognition of shape is based on information adjustment to mental patterns rather than on depicting a perceived entity pixel by pixel. In that case, our experience of a particular shape, for example the shape of a chair, always starts when the amount of perceived information is enough to assure our brain that what we see is most probably a chair rather than an elephant. Obviously, the recognition of everything, not only a chair, is influenced by the whole context of a perception (Stein, Wallace, & Stanford, 2001): what we saw earlier, what we see apart from a chair, what we hear or smell, what has been said to us, etc. Nonetheless, the general rule is that our brains operate using a set of mental patterns to create pictures of the environment that we perceive.

An interesting alternative for the cognitivist accounts of perception is the idea of inactive cognition (Matyja & Schiavio, 2013). According to the advocates of this view, our perception is a result of the dynamic interactions between our body and the environment (Varela, Thompson, & Rosch, 1991) rather than a unidirectional process of creating mental representations. In other words, perception is not merely the recognition of pre-existing things or events, but also a dynamic, ongoing process (Reybrouck, 2005). From this perspective, hearing and action in response to what is heard are inextricably connected. In other words, what is heard depends on what the body can and what it cannot do, in terms of Gibson’s idea of affordances (Gibson, 1983). However, independent of which point of view one prefers—inactive perception or perception by means of representations—audition cannot be seen as a mere form of reflecting reality. Nevertheless, there are some perceptual phenomena that seem to be much more

detached from the perceived reality than others. A good example of such a phenomenon is the perception of speech that depends on social learning to a greater degree than, for instance, the simple recognition of a chair in our environment. Auditory perception seems to be particularly susceptible not only to the species-specific but also to the culture-specific interpretation of stimuli. After all, the recognition of phonemes in a foreign language is certainly a much more difficult task in comparison to the recognition of an object's shape. This difference is most probably due to the important role that auditory perception plays in human communication (Wiley, 2015), since the auditory channel not only became the default modality of language but also serves a crucial function in the exchange of many other socially valid cues, for example by means of speech prosody, which is based to some extent on social learning. Moreover, the auditory channel has become a default mode not only for speech and singing but also for other human-specific forms of vocal communication such as crying and laughing, (Hauser, 1996). Although humans are able to use other channels in order to intentionally transmit information (e.g., the visual channel in the case of gestures or facial expressions, etc.), the complexity of speech and music (which allows us to communicate an enormous scope of knowledge) as well as the ease of their recognition suggest that the human auditory channel has evolved not only as a tool of hearing environmental sounds but also as a tool whose function is strictly connected with the communication of specific information (Fitch & Jarvis, 2013). In other words, the spectral and temporal characteristics of sound has not only become a source of information about the external world, but has also started to serve as a medium of communication (Reybrouck & Podlipniak, 2019). As a result, the manipulation of acoustic features by the sender transforms sound into an intentional signal. In order to fulfil this function, sound has become a useful means to trigger similar mental patterns (i.e., the patterns of brain activity) in the brains of communicating individuals. Unlike the mental patterns of the perceived physical objects however, these patterns are accessible only to the individuals who have acquired a culture-specific code. Since human culture is a property of human brains and not of the external physical environment, then the patterns of this code must, in fact, be related to mental phenomena rather than to external physical objects. From this point of view, every musical and speech sound is a means by which a culture-specific message is encrypted in the statistical distribution of acoustic properties.

Of course, this does not mean that music and speech are only composed of features that convey a culture-specific message. Every music and speech stimulus is a source of an enormous number of clues which are inferred by

listeners during perception, and also by the means of evolutionarily old mechanisms independent of cultural influence, such as the location of the sound source or the recognition of the size of the sound source (Bregman, 1990).

However, what differentiates music and speech from environmental sounds, such as traffic noise or wolves howling, is their perceivable structure organized according to specific rules. Without the knowledge of the rules that govern a particular medium of communication, regardless of whether it is music or speech, the information coded in this medium is inaccessible, as in the case of a foreigner listening to an unfamiliar language. However, the already mentioned statistical distribution of the acoustic properties of sounds (e.g., the formants of particular vowels) is enough to initiate the interpretation of these perceived sounds in terms of specific structural patterns (e.g., particular phonemes) in the brains of the individuals who possess tacit knowledge about a particular language. In the case of language, these patterns represent phonemes, morphemes, words, etc. which are composed according to phonotactical and grammatical rules. All these elements together constitute the structure of speech that one always experiences when listening to speech.

WHAT IS MUSICAL STRUCTURE?

Musical structure is often compared to speech, as its elements can be recognized as discrete units organized according to tacit rules, just like in speech (Lerdahl, 2013). From the psychological perspective the musical structure of tonal music is composed of pitch classes (pitch chroma) (Shepard, 1982) distributed in time (London, 2012). In musical analysis this structure can be described in terms of pitch intervals and rhythm measures (e.g., Hed, Gjerdingen, & Levin, 2015). Both these elements are discrete. There are also a restricted number of pitch intervals and rhythm measures in all known natural musical systems. From this point of view, both musical and speech structures represent what is known as the Humboldt system (Merker, 2002)—i.e., a system that is composed of a finite set of distinct elements combined in a generative way.

However, while phonemes in speech are unquestionably treated as solely mental phenomena (Berent, 2013), musical intervals and rhythm measures are usually described in terms of physical properties (Roederer, 2008). Even tonal relations between pitch classes in melody are often claimed to result from the relations between the frequencies of harmonics of consecutive sounds. From this perspective, every melody is an acoustic entity that can be recognized by means

of acoustic analysis, no matter whether this analysis is conducted by the human brain, a bird's brain, or a computer. The belief about the acoustic nature of musical structure still seems to dominate in musicology and the philosophy of music, despite the fact that many observations render this belief untenable.

First of all, the perception of musical pitch in a musical context is not very sensitive to frequency deviations, for instance from equal temperament, although our perception of pitch in general is very precise in terms of tuning (Rakowski, 2009). In other words, there is a zone of sound frequencies in which all sounds are interpreted by our nervous system as belonging to the same pitch class. However, the interpretation of two consecutive sounds as a particular pitch interval depends on musical culture. Secondly, there is a huge variety of musical systems around the world, which means that the same sound stimulus can be recognized as a different musical interval by two individuals nurtured in different cultures. Thirdly, even a seemingly nonmusical stimulus such as speech can be heard as melody under some conditions (Deutsch, Henthorn, & Lapidis, 2011; Deutsch, Lapidis, & Henthorn, 2008). The same is true in the domain of rhythm. The recognition of prominence in meter (metrical weight) and musical pulse are in fact also an imposition of internally experienced periodicity on perceived sounds (London, 2012). Moreover, the recognition of every melody as a single entity is in fact unexplainable solely in terms of acoustics similar to the recognition of a sentence in speech. Our feelings of completion and suspense in response to sounds heard in a melody cannot be inferred from the acoustic properties of these sounds (Krumhansl & Cuddy, 2010). After all, the acoustically same sound can be experienced as suspense or completion depending on the pitch context. All these musical traits—i.e., pitch intervals, rhythm measures, metrical weight, musical pulse, and tonal relations—are the crucial elements of almost every musical structure. They are depicted in contemporary standardized musical notation and are the subjects of many musicological analyses.

Although music as a whole can be understood as “a collection of sound/time phenomena” (Reybrouck, 2005, p. 235), the perception of sound while listening to music leads to a complex processing of information which gives new psychological qualities to the perceived sounds. This processing is hierarchical, and its hierarchy reflects the evolutionary age of cognitive mechanisms involved in this process. The lower level of processing is based on evolutionarily older mechanisms compared to the higher cognitive level of processing that is based on evolutionarily younger mechanisms (Reybrouck, 2005). Since the processing of musical structure is a human-specific trait which depends on social learning, it must include higher-level processing.

Taking into account all of the above-mentioned properties of music perception, the claim that musical structure resides in the acoustic properties of sound is unconvincing (Podlipniak, 2017b). Bearing this in mind and as far as we currently understand, every musical structure actually exists in the human brain as specific patterns of bioelectrical activities. In the case of music perception, the particular acoustic properties of sound (e.g., F_0) act as a medium that triggers similar musical patterns in the brains of communicating individuals (Bharucha, Curtis, & Paroo, 2011). These patterns are a result of music-specific interpretation of sound stimuli in the context of all other information detected by the other human senses that can also influence this interpretation.

Although sound can be detected not only by the auditory channel but also by the vestibular (Todd & Cody, 2000; Todd, Cody, & Banks, 2000) and tactile systems (Huang, Gamble, Sarnlertsophon, Wang, & Hsiao, 2012), the most important acoustic cues that allow us to experience a musical structure are detected by the organ of Corti, located in the inner ear (Roederer, 2008). This, however does not mean that information obtained by the vestibular or tactile systems cannot contribute to our experience of musical structure. It is possible that the sensation of a musical rhythm can be emphasized, for instance, by the detection of low frequency sounds by our sense of touch (Huang, Gamble, Sarnlertsophon, Wang, & Hsiao, 2012). Also, the sensation of sound periodicity can be emphasized by the vestibular and tactile systems, which can facilitate our experience of musical meter. Since both rhythm and meter are the crucial elements of the metro-rhythmical structure of music (Lerdahl & Jackendoff, 1983), one cannot ignore the possible contribution of both the vestibular (Trainor, Gao, Lei, Lehtovaara, & Harris, 2009) and tactile senses to the construction of our mental structural patterns while experiencing music. Nonetheless, it is less probable that vestibular or tactile sensations would be able to deliver sufficient cues for the recognition of a musical pitch structure. It is also possible that the experience of the metro-rhythmical structure of music is influenced by the sensation of body movements (Levitin, Grahn, & London, 2018). In fact, the metro-rhythmical structure of music seems to be multimodal (Su & Salazar-López, 2016) and can be inferred from dance movements. For the recognition of pitch intervals and, in consequence, also melodies, the activity of the organ of Corti seems to be indispensable. Only this organ is able to analyze the spectral and temporal features of sound precisely to an extent sufficient to allow the anatomical structures of our auditory pathway to generate the dynamic cognitive representations of pitch structure.

The process of generating auditory representations is often called “auditory streaming” (Bregman, 1990; McAdams & Bregman, 2014) because it resembles the constant flow of water in a stream. An interesting alternative conception of auditory perception that extends the sound experience to embodied and enacted cognition is the idea of ecological perception and event perception (Matyja, 2016; Matyja & Schiavio, 2013; Reybrouck, 2005; Schiavio & van der Schyff, 2016). From both perspectives, however, the representations of every musical structure differ from the representations of environmental sounds in terms of reference. While the representations of environmental sounds are used by our brains to recognize particular objects and their properties (e.g., their size, shape, velocity, as well as gender, attractiveness, etc. in the case of the human voice), the representations of musical structure seem to be a source of a conspecific and even culture-specific message. In other words, the representations of environmental sounds refer to external objects such as a falling tree, a flying airplane, or uncle Bob. The representations of musical structure refer to the intentions of singing (or playing) individuals. These intentions are, of course, not necessarily conscious. If musical structure is a result of human-specific proclivities to organize and interpret sounds in terms of pitch intervals organized in psychological time (Podlipniak, 2016, 2017a), then the message of every musical structure must depend on the biological function of these human-specific proclivities.

The intention behind a musical structure is, of course, not the only message that one can infer from perceived sounds whilst listening to music. A music stimulus is usually a complex acoustic event that consists of many acoustic features interpreted by our nervous system according to many patterns related to different sorts of information (Hauser, 2000; Panksepp & Bernatzky, 2002; Schiavio & van der Schyff, 2016; Zimmermann, Leliveld, & Schehka, 2013) which are not parts of musical structure. For instance, some of these patterns are part of an evolutionarily ancient form of vocal communication shared by humans with other mammalian species and known as “expressive dynamics” (Merker, 2003). Although it is intuitively used by humans in music, this form of vocal communication is not music-specific, as it is also present in speech. Similarly, the recognition of timbre is based on the abilities used in the perception of auditory phenomena other than music—for example, the recognition of an animal based on its vocalization such as dog barking, the recognition of a particular person based on his voice timbre, etc. Despite the fact that the experience of timbre in music is often a very important part of our mental picture of a musical piece, not only in some Western music genera but also in certain extra-European musical cultures (Nikolsky, 2015), the timbral characteristics of music seem to be something

additional. After all, a melody played on different musical instruments or sung by different singers is recognized as the same melody. However, the actual reasons for this diminished role of musical timbre in the mental prototypes of musical structure, whether it is influenced by cultural factors or just due to pitch and rhythm dominance in our biological specificity of musical communication, is a matter of hot debate.

THE PERCEPTION OF MUSIC OR THE MUSICAL PERCEPTION OF SOUNDS?

Even the phrase “perception of music” suggests that music is something external to the subject perceiving it. This assumption leads to the conclusion that there must be something like a “musical sound” which differs from other types of sound. The problem of the musical properties of sound had been a subject of debate within musicology (Handschin, 1948) right up until the twentieth century, with the wide acceptance of avant-garde music as actually being music. Yet, the dependence of the definition of music on an arbitrary choice of composers (or listeners) is, to say the least, unsatisfactory from the perspective of music psychology and biomusicology (Brown, Merker, & Wallin, 2000). If music perception is an objective process that differs from the perception of speech and other sound phenomena, then there must be some objective criteria which allow us to determine what the specific traits of music are. The question is, however, whether these traits are acoustic or psychological. In the above-mentioned traditional musicological debate on the musicality of sound, the often indicated musical trait of sound was its harmonicity that, according to neo-Pythagoreans, leads to specific musical relations between perceived musical pitches. Nevertheless, these same harmonic sounds are always used in speech when we are pronouncing vowels. In this case, however, they are perceived not as music but as speech. Therefore, another important condition has been added in order to differentiate musical sounds from speech sounds—the relative stability of the fundamental frequency of a harmonic sound (F_0) in time. According to this view, the main difference between speech and singing is that sounds in singing are organized around some “frequency guideposts,” which results in greater intonation accuracy in singing than in speech (Zatorre & Baum, 2012). However, as mentioned above, in some conditions even speech sounds, the intonation of which is not stable according to music criteria, can be recognized as a melody (Deutsch et al., 2011; Deutsch et al., 2008). Moreover, the pitches of this melody are perceived

hierarchically in terms of tonal relations, which means that some pitches are more important than others: for instance, C in C major (tonic) is perceived as more important (stable, relaxing) than C sharp (an out-of-key note), which is less important (unstable, tense). Yet this tonal hierarchy is absent from our experience of the acoustically same sounds before the perceptive transformation of spoken words into melody in the “speech to song” illusion. A similar musical experience of perceived sounds is present when we are listening to songbirds’ songs, which are in fact forms of sound communication specific to another species and are based on birds’ mental patterns that are different from human mental patterns (Shannon, 2016). Therefore, the most promising explanation of music perception specificity is the assumption that the objective criteria of music are the structural features which are in fact our mental categories based partly on human-specific inherited proclivities and partly on learning. As a result, our brains are sensitive to certain acoustic features. If there are enough of them to persuade us that what we hear is music, our brains start to interpret these stimuli as musical. From this perspective, it is more precise to say that we perceive sounds in a musical way rather than that we perceive music.

CONCLUSION

The neo-Pythagorean view of musical structure is not only present in “folk psychology,” but is still also very popular in musicology and related disciplines. As has been emphasized, psychological studies show that musical structure does not exist as an acoustic musical trait but is a property of the human mind. From this perspective, it is not only pitch intervals that must be understood as psycho-cultural phenomena (Parncutt & Hair, 2018) but also the other structural features mentioned above, such as metrical grid, tonal relations, and rhythm measures. However, not all aspects of our musical experience related to musical structure are fully understood. The most intriguing issue related to the experience of musical structure is the role of timbre in the structural categories of certain musical styles observed in non-European cultures. Also the question about conditions on which some “non-musical” sound stimuli are perceived as music is waiting for an answer. All these problems necessitate further research. The answers to these questions would be very helpful in the search for the nature of neural specialization of music processing as well as the origin of music.

REFERENCES

- Berent, I. (2013). The phonological mind. *Trends in Cognitive Sciences*, 17(7), 319–327. <https://doi.org/10.1016/j.tics.2013.05.004>
- Bharucha, J., Curtis, M., & Paroo, K. (2011). Musical communication as alignment of brain states. In P. Rebuschat, M. Rohrmeier, J. A. Hawkins, & I. Cross (Eds.), *Language and music as cognitive systems* (pp. 139–155). Oxford, New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199553426.003.0016>
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, London: The MIT Press.
- Brown, S., Merker, B., & Wallin, N. L. (2000). An introduction to evolutionary musicology. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 3–24). Cambridge, London: The MIT Press.
- Caleon, I., & Ramanathan, S. (2008). From music to physics: The undervalued legacy of Pythagoras. *Science & Education*, 17(4), 449–456. <https://doi.org/10.1007/s11191-007-9090-x>
- Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception: An Interdisciplinary Journal*, 5(2), 197–205. <https://doi.org/10.2307/40285392>
- Crocker, R. L. (1964). Pythagorean mathematics and music. *The Journal of Aesthetics and Art Criticism*, 22(3), 325. <https://doi.org/10.2307/427236>
- 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. <https://doi.org/10.1121/1.3562174>
- Deutsch, D., Lapidis, R., & Henthorn, T. (2008). The speech-to-song illusion. *The Journal of the Acoustical Society of America*, 124(4), 2471. <https://doi.org/10.1121/1.4808987>
- 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). Cambridge, London: The MIT Press.
- Gibson, J. J. (1983). *The senses considered as perceptual systems*. Westport, CT, US: Greenwood Press.
- Handschin, J. (1948). *Der Toncharakter: eine Einführung in die Tonpsychologie* [The sound character: An introduction to the psychology of sound]. Zürich, Swiss: Atlantis.
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge, London: The MIT Press.
- Hauser, M. D. (2000). The sound and the fury: Primate vocalizations as reflections of emotion and thought. In N. L. Wallin, S. Brown, & B. Merker (Eds.), *The origins of music* (pp. 77–102). Cambridge, MA, US: The MIT Press.
- Hed, S., Gjerdingen, R. O., & Levin, D. (2015). Pitch-and-rhythm interrelationships and musical patterns: Analysis and modelling by subdivision schemes. *Journal of Mathematics and Music*, 9(1), 45–73. <https://doi.org/10.1080/17459737.2014.980344>
- Helmholtz, H. von (1954). *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Trans.) (2nd ed.). New York, NY, US: Dover Publications.
- Huang, J., Gamble, D., Sarnlertsophon, K., Wang, X., & Hsiao, S. (2012). Feeling music: Integration of auditory and tactile inputs in musical meter perception. *PLoS ONE*, 7(10), e48496. <https://doi.org/10.1371/journal.pone.0048496>
- Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, London: The MIT Press.
- Kivy, P. (1990). *Music alone: Philosophical reflections on the purely musical experience*. Ithaca, London: Cornell University Press.

- 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). New York, Dordrecht, Heidelberg, London: Springer New York. https://doi.org/10.1007/978-1-4419-6114-3_3
- Large, E. W., & Almonte, F. V. (2012). Neurodynamics, tonality, and the auditory brainstem response. *Annals of the New York Academy of Sciences*, 1252(1), E1–E7. <https://doi.org/10.1111/j.1749-6632.2012.06594.x>
- Large, E. W., & Tretakis, A. E. (2005). Tonality and nonlinear resonance. *Annals of the New York Academy of Sciences*, 1060(1), 53–56. <https://doi.org/10.1196/annals.1360.046>
- Lerdahl, F. (2013). Musical syntax and its relation to linguistic syntax. In M. A. Arbib (Ed.), *Language, music, and the brain* (pp. 257–272). Cambridge, London: The MIT Press. <https://doi.org/10.7551/mitpress/9780262018104.003.0010>
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, London: The MIT Press.
- Levitin, D. J., Grahn, J. A., & London, J. (2018). The psychology of music: Rhythm and movement. *Annual Review of Psychology*, 69(1), 51–75. <https://doi.org/10.1146/annurev-psych-122216-011740>
- London, J. (2012). Three things linguists need to know about rhythm and time in music. *Empirical Musicology Review*, 7(1–2), 5–11.
- Matyja, J. R. (2016). Embodied music cognition: Trouble ahead, trouble behind. *Frontiers in Psychology*, 7, 1891. <https://doi.org/10.1086/392866>
- Matyja, J. R., & Schiavio, A. (2013). Enactive music cognition: Background and research themes. *Constructivist Foundations*, 8(3), 351–357.
- McAdams, S., & Bregman, A. (2014). Musical streams. *Computer Music Journal*, 3(4), 26–43.
- Merker, B. (2002). Music: The missing Humboldt system. *Musicae Scientiae*, 6, 3–21. <https://doi.org/10.1177/102986490200600101>
- Merker, B. (2003). Is there a biology of music? And why does it matter? In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *Proceedings of the 5th Triennial ESCOM Conference* (pp. 402–405). Hanover: Hanover University of Music and Drama.
- Nikolsky, A. (2015). Evolution of tonal organization in music mirrors symbolic representation of perceptual reality. Part-1: Prehistoric. *Frontiers in Psychology*, 6, 1405. <https://doi.org/10.3389/fpsyg.2015.01405>
- Nolan, C. (2002). Music theory and mathematics. In T. Christensen (Ed.), *The Cambridge history of Western music theory* (pp. 272–304). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CHOL9780521623711.012>
- Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: The neuro-affective foundations of musical appreciation. *Behavioural Processes*, 60(2), 133–155. [https://doi.org/10.1016/S0376-6357\(02\)00080-3](https://doi.org/10.1016/S0376-6357(02)00080-3)
- Papadopoulos, A. (2002). Mathematics and music theory: From Pythagoras to Rameau. *The Mathematical Intelligencer*, 24(1), 65–73. <https://doi.org/10.1007/BF03025314>
- Parncutt, R., & Hair, G. (2018). A psychocultural theory of musical interval: Bye bye Pythagoras. *Music Perception: An Interdisciplinary Journal*, 35(4), 475–501. <https://doi.org/10.1525/mp.2018.35.4.475>
- Podlipniak, P. (2016). The evolutionary origin of pitch centre recognition. *Psychology of Music*, 44(3), 527–543. <https://doi.org/10.1177/0305735615577249>
- Podlipniak, P. (2017a). The role of the Baldwin effect in the evolution of human musicality. *Frontiers in Neuroscience*, 11, 542. <https://doi.org/10.3389/fnins.2017.00542>
- Podlipniak, P. (2017b). Tonal qualia and the evolution of music. *Avant*, 8(1), 33–44. <https://doi.org/10.26913/80102017.0101.0002>

- Raftopoulos, A. (2009). *Cognition and perception: How do psychology and neural science inform philosophy?* Cambridge, MA, US: The MIT Press.
- Rakowski, A. (2009). The domain of pitch in music. *Archives of Acoustics*, 34(4), 429–443.
- Reybrouck, M. (2005). A biosemiotic and ecological approach to music cognition: Event perception between auditory listening and cognitive economy. *Axiomathes*, 15(2), 229–266. <https://doi.org/10.1007/s10516-004-6679-4>
- Reybrouck, M., & Podlipniak, P. (2019). Preconceptual spectral and temporal cues as a source of meaning in speech and music. *Brain Sciences*, 9(3), 53. <https://doi.org/10.3390/brainsci9030053>
- Roederer, J. G. (2008). *The physics and psychophysics of music: An introduction*. New York, NY, US: Springer.
- Schiavio, A., & van der Schyff, D. (2016). Beyond musical qualia. Reflecting on the concept of experience. *Psychomusicology: Music, mind, and brain*, 26(4), 366–378. <https://doi.org/10.1037/pmu0000165>
- Scruton, R. (1999). *The aesthetics of music*. Oxford, New York: Oxford University Press. <https://doi.org/10.1093/019816727X.001.0001>
- Sethares, W. A. (1998). *Tuning, timbre, spectrum, scale*. London: Springer. <https://doi.org/10.1007/978-1-4471-4177-8>
- Shannon, R. V. (2016). Is birdsong more like speech or music? *Trends in Cognitive Sciences*, 20(4), 245–247. <https://doi.org/10.1016/j.tics.2016.02.004>
- Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *Psychology of music* (1st ed., pp. 343–390). New York, NY, US: Academic Press. <https://doi.org/10.1016/B978-0-12-213562-0.50015-2>
- Stein, B. E., Wallace, M. T., & Stanford, T. R. (2001). Brain mechanisms for synthesizing information from different sensory modalities. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 710–736). Malden, MA, Oxford: Blackwell.
- Su, Y.-H., & Salazar-López, E. (2016). Visual timing of structured dance movements resembles auditory rhythm perception. *Neural Plasticity*, 2016, Article ID 1678390. <https://doi.org/10.1155/2016/1678390>
- Todd, N. P. M., & Cody, F. W. (2000). Vestibular responses to loud dance music: A physiological basis of the “rock and roll threshold”? *The Journal of the Acoustical Society of America*, 107(1), 496–500.
- Todd, N. P. M., Cody, F. W., & Banks, J. R. (2000). A saccular origin of frequency tuning in myogenic vestibular evoked potentials?: Implications for human responses to loud sounds. *Hearing Research*, 141(1–2), 180–188.
- Trainor, L. J., Gao, X., Lei, J., Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex*, 45(1), 35–43. <https://doi.org/10.1016/J.CORTEX.2007.10.014>
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA, US, London: The MIT Press.
- Wiley, R. H. (2015). *Noise matters: The evolution of communication*. Cambridge, London: Harvard University Press.
- Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, 10(7), Article ID e1001372. <https://doi.org/10.1371/journal.pbio.1001372>
- Zimmermann, E., Leliveld, L., & Schehka, S. (2013). Toward the evolutionary roots of affective prosody in human acoustic communication: A comparative approach to mammalian voices. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *Evolution of emotional communication: From sounds in nonhuman mammals to speech and music in man* (pp. 116–132). Oxford, New York: Oxford University Press.