



# Hearing chimeras

Elvira Di Bona<sup>1</sup>

Received: 29 June 2021 / Accepted: 27 April 2022  
© The Author(s) 2022, corrected publication 2022

## Abstract

I argue that chimericity is a property that we typically experience when listening to multi-instrumental music. It is the property of hearing as a unified whole a melody or a harmony that does not belong to any single sound source but instead consists of the assembling of melodic or harmonic fragments coming from different sources. Chimericity is not reducible to the low-level audible properties of pitch and loudness; it is cognized at the perceptual level thanks to the auditory mechanism of primitive grouping. My aim is to show that chimericity is a perceptual property, one that we can genuinely hear. In developing this view, I engage with the debate within the philosophy of sense perception between rich vs thin views of the content of perceptual experience.

**Keywords** Perception · Musical experience · Auditory experience · Auditory streams · Musical streams · Rich view · Gestalt principles

## 1 Introduction

I argue that chimericity is a grouping property that we can perceptually experience when, in the course of listening to multi-instrumental music—music played by more than one instrument—we hear as a whole a melody or a harmony that does not belong to any single sound source, meaning that it has not been produced by a single sound source but is instead the result of the assembling of melodic or harmonic fragments coming from different sound sources. I take the concept of a chimera from Greek mythology. The chimaera was a beast made of parts belonging to different animals: the head of a lion, the body of a goat, and the tail of a dragon. I use the term metaphorically to refer to an auditory compound that does not belong to any single sound source, in the sense that it is not produced by a single musical instrument but is instead the result of a combination of auditory fragments deriving from different sound sources. Examples of melodic chimeras are the melodic lines that we hear in many of Antonio Vivaldi's

---

✉ Elvira Di Bona  
elvira.dibona@unito.it

<sup>1</sup> Department of Philosophy and Education Sciences, University of Turin, via S. Ottavio 20, 10124 Turin, Italy

concertos for two, three, and four instruments. In these cases, we hear melodic lines that are unified wholes, even though they have been produced by different soloists playing different instruments, with each musician responsible for a specific melodic fragment. We quite often hear this melodic line as a temporally extended whole, as if a single musical source had produced it. As we experience melodic chimeras, we also experience harmonic chimeras, for example, when we hear unitary wholes of simultaneous sounds as constituting chords, even though the sounds of the chords have been produced by different musical instruments. Harmonic chimeras are extremely common in music, and they are widespread in Western musical traditions, from classical to jazz, pop, and rock. Just to name one example, one can think of the opening chords of *The Young Persons' Guide to the Orchestra* (op. 34) by Benjamin Britten, in which the entire orchestra plays sounds together and thereby produces a chord progression.

Auditory chimeras are widespread in multi-instrumental music and are usually created by the intertwining of two or more musical voices or instruments. The experience of musical chimeras contributes to the appreciation of music since it allows the listener to grasp some of the essential musical features of a vast majority of musical pieces, namely specific melodies and harmonies that constitute the basic elements of musical expressivity.

For instance, in Anton Webern's orchestration of the *Ricercar* from "The Musical Offering" by Johann Sebastian Bach, the theme of the fugue that opens the composition, which is a crucial musical element, is elaborated sequentially by different wind instruments, generating a melodic chimera. Even if we perceive a subtle change of timbre, one that confers to the melody a specific nuanced allure, we perceive the theme of the fugue—conveyed in its entirety by different musical instruments—as it should be, namely, as a unitary melody. Analogously, the simultaneous sounds emitted by the entire orchestra of strings and winds in the opening three bars of the overture of Wolfgang Amadeus Mozart's *Die Zauberflöte*, which form a harmonic chimera, are perceived as they are intended to be, namely, as unitary chords. This is an essential musical feature of the overture, which anticipates the further melodic development carried out by the strings and the bassoon.

Being able to detect and individuate melodies and harmonies when listening to any genre of music (pop, rock, jazz or classical) is an important ability for the appreciation of music, since it allows one to be acquainted with the general structure and the different aspects of a musical piece. We do not necessarily need to be able to experience the fact that the melodies and harmonies we detect are generated by different instruments. That is, we do not need to be aware of actually *hearing* auditory chimeras; we can also just as easily be under the illusion that it is just one single instrument that generates them. What is essential for the experience of musical chimeras is that we track a melody or a harmony by virtue of experiencing a feeling of unity, even though they are generated by different sound sources.

To be more precise, the unity we experience is a "composite" unity.<sup>1</sup> That is, we have the feeling that, despite the fact that the melodic and harmonic fragments coming from different sources are unified in a single melody and a single harmony, these fragments

---

<sup>1</sup> Throughout the paper, unless otherwise specified, instead of using the expression "feeling of composite unity" when talking about multi-instrumental music, I use the simpler expression "feeling of unity."

remain distinguishable. This is due to the timbral differences between the diverse musical instruments producing the musical fragments. Similarly, it is the feeling of composite unity that guarantees that, when we look at the Greek Chimera, we see a single animal, and yet, we can also tell that that single animal is made up of parts identifiable as belonging to different animals. The feeling of composite unity reveals the presence of chimericity because its content is constituted by the (perceivable) property of chimericity itself. Needless to say, that chimericity is *perceivable* has to be shown, and this is precisely the aim of this paper. I will argue that the property of chimericity is a perceptual property, one that we can genuinely hear. I engage with the debate within the philosophy of sense perception between rich vs thin views of the content of perceptual experience. My conclusion is that, by virtue of a method that proceeds by analogy, we can embrace a rich view within musical perception as well, at least when focusing on the property of chimericity.<sup>2</sup>

This paper is structured as follows: in Sect. 1, I introduce the distinction between the rich and the thin view of perception, and I briefly present, in Sect. 2, the different methods employed in the recent literature on the content of perception to show that a specific property is perceivable. In the same section, I also present my own methodology, which is a form of reasoning that proceeds by analogy. Then, I fully elaborate my method by describing, in Sect. 3, how the auditory system employs a perceptual mechanism to segregate and individuate sound streams when exposed to environmental sounds and, then, in Sect. 4, by showing how the same mechanism is at work when listening to musical sounds and segregating them into musical streams generating auditory chimeras. In Sect. 5, I address some worries that might challenge my view. Finally, in Sect. 6, I conclude that my methodology leads to the claim that chimericity is an audible property that we are capable of perceiving when listening

---

<sup>2</sup> I do not exclude that melodies or harmonies played by a single instrument could contribute to the same point but: (1) I wanted to focus on a specific type of melody in multi-instrumental music, which is the kind of music to which we are generally exposed. Chimericity is important for the appreciation of this kind of music. (2) Also, it has to be highlighted that the sense of unity arises despite the fact that the melodic fragments come from different instruments. This emphasizes the relevance of speaking in terms of chimericity rather than a solo instrument melody. In the case of solo instrumental music, the sense of unity is obvious, and there is no need to highlight it, since a solo instrument plays a melody unitarily by default. Moreover, this sense of unity is not “composite”—as it is, instead, in the case of multi-instrumental music—since a single instrument, having just one timbre, generates simple unified music or harmonic wholes. This means that solo instrumental music does not generate chimeras. Even though I do not exclude that single-instrument music could be used in favor of the rich view by appealing to the sense of unity as revealing melodies and harmonies, this view cannot be justified by appealing to chimeras; rather, it should be justified on different grounds, which I leave for future research. (3) In the paper, I am not talking only about melodic chimeras but also about harmonic chimeras. Even though we can experience a sense of unity when listening to harmonies produced by solo instruments, this is not feasible on many individual musical instruments, given that the majority of instruments are monodic (except, of course, for instruments such as the piano, accordion, harpsichord, organ, harp, guitar, and a few more).

to music.<sup>3</sup> Therefore, we have reasons sufficient to support a rich view of auditory experience when focusing on listening to musical sounds.<sup>4</sup>

## 2 Rich view vs thin view

We can group views in the philosophy of perception about the content of perceptual experience into two camps—rich (thick) views and thin (sparse) views. Those who support the thin or sparse view argue that we only perceptually represent the low-level properties of color, shape, and size in vision or the properties of pitch and loudness in audition (Tye, 1995; Dretske, 1995; Clark, 2000). Those who support the rich or thick view argue that we also perceptually represent properties usually labeled high-level properties, such as being a pine tree, a face expressing certain emotions, the relation of causation that connects different events, being an edible object, the meaning articulated by a given speech, or the gendered property of a human voice (Siegel, 2009, 2010; Butterfill, 2009; Block, 2014; Nanay, 2011, 2012; Peacocke, 1992; Strawson, 1994/2010; Di Bona, 2017). The diversity of methods (Fish, 2013; Helton, 2016; Masrour, 2011) employed in the recent literature on the topic seems to suggest that we cannot tell a priori which is the best way to establish whether a property (or a group of properties) is perceivable. The only way to proceed is to focus on the property (or the group of properties) under examination and to find out what would be a good method to use in order to test whether this specific property is perceivable.

To evaluate whether chimericity is audible, I discuss empirical results by virtue of reasoning by analogy. My strategy is based on the assumption that, with respect to environmental sounds or in ordinary auditory contexts (i.e. when we hear sounds that commonly surround us in everyday life, such as the sound of a door slamming, the jiggling of a key, the breaking of a glass, or people laughing), there is a perceptual mechanism called primitive grouping (Bregman, 1990) that makes possible the segregation and the grouping of sounds, thereby generating the perceptual properties of being a kind of environmental sound.

This perceptual mechanism makes possible the grouping and the segregation of sounds and the distinction between one sound from another. Specifically, it makes it possible to distinguish between a sound belonging to the category of musical sounds

---

<sup>3</sup> In this paper, I will expand on Bregman's idea of "chimeric assignment" (1990, p. 450), consisting in the event in which we "hea[r] a sentence that was created by the accidental composition of the voices of two persons who just happens to be speaking at the same time". Whereas Bregman only mentions harmonic chimeras briefly, I discuss both harmonic and melodic chimeras at length because my aim is to introduce the concept of chimericity within the debate around the admissible content of perception.

<sup>4</sup> I can be neutral on the metaphysics of sound since my point on the perception of chimericity stands whatever sounds *are*. In fact, the metaphysics of sound merely influences the space where we "locate" the musical chimera that we hear. Indeed, if sounds are distal events identical to the vibration of the sounding object (Casati & Dokic, 1994; Casati et al., 2013; Pasnau, 2009), relational events heard as being located where the sounding objects are (O'Callaghan, 2007, 2009), stable properties of the sounding object (Pasnau, 1999) or dispositional properties that objects manifest when appropriately stimulated (Kulvicki, 2008), then we perceive a chimera as a unitary whole made of sounds taking place in the proximity of the sounding object (or the source of sound). If sounds are either identical to sound waves (O'Shaughnessy, 2000; Perkins, 1983; Sorensen, 2009) or not identical to sound waves but instantiated by them (Nudds, 2009), then we perceive a musical chimera as a unitary whole made of sounds taking place in the medium.

(i.e. sounds produced by an orchestra, by a tenor rehearsing, or by street musicians) and a sound belonging to the category of speech sounds (i.e. ones that are produced by someone speaking and which have a semantic meaning). However, the perceptual mechanism is not responsible for the full recognition of the sounds we are listening to, since to fulfill this specific task the auditory system employs a higher-level mechanism that is known as schema-based grouping. The assumption is that if, within an ordinary auditory context, this perceptual mechanism generates perceptual properties that are the properties of being an environmental sound, we have good reasons to think that if, in a different auditory context, such as a musical one, the same perceptual mechanism were to take place, it would likely generate perceptual properties as well. In other words, the method consists in determining whether the mechanism that segregates and groups sound streams when listening to environmental sounds, which is a perceptual mechanism—a mechanism that allows us to hear the perceptual property of being a kind of environmental sound—is also at work when we have to segregate and group streams of musical sounds as to form auditory chimeras. If the perceptual mechanism giving rise to perceptual properties—such as the distinct auditory sound streams we segregate when hearing environmental sounds—is also at work in the musical scene analysis that gives rise to auditory chimeras, then we can conclude that such properties are indeed perceptual properties as I maintain them to be.

### 3 The auditory scene analysis and environmental sounds

Bregman (1990) deals with the mechanisms at the basis of the grouping, segregation, and recognition of the auditory streams that inhabit the auditory landscape. The auditory scene analysis begins when the auditory system faces up to the chaotic *multiplicity* of auditory stimuli coming from different directions and generated by different sources, organizing them as constituting meaningful auditory streams. This analysis is based on a two-passage mechanism: primitive grouping and schema-based grouping. Primitive grouping (ibid: chapters II and III) is a perceptual process that accomplishes the task of first detecting auditory stimuli and then segregating them into auditory streams, so that we can distinguish between different kinds of sounds or streams of sounds. Schema-based grouping, on the other hand, is responsible for the identification and categorization of streams by virtue of the application of conceptual schemas stored in the memory. I focus only on primitive grouping since my aim is to determine whether the perceptual grouping mechanism responsible for the detection and segregation of environmental sounds works analogously when grouping auditory streams and musical streams forming auditory chimeras. Primitive grouping begins by decomposing the amalgam that constitutes the auditory *multiplicity* into elementary auditory elements, such as various loudnesses, pitches, and timbres, which are the specific audible qualities we commonly attribute to sounds. After the detection of these auditory stimuli, the challenge of the auditory scene analysis at the primitive grouping level is to understand “what” groups with “what”. That is to say that the auditory system has to figure out how to correctly group together the sensory elements belonging to the environmental source events that veridically generated them. It succeeds with this task thanks to a process that works on two levels: *sequential* and *simultaneous*. The

auditory qualities are grouped into streams via sequential integration when there is a sequence of sounds with different speed and pitch that come in a series and need to be put together as they evolve over time. The auditory qualities are grouped, instead, into streams via the simultaneous integration that combines sounds of different pitch occurring at the same time, and they are thereby fused into a unified stream. The two processes operate by the employment of certain principles, some of which are Gestalt principles applied to audition (Bregman, 1990, pp. 196–202, 248–293). The Gestalt principles are *proximity*, according to which we tend to group together sounds that are near to one another in time, *frequency* (which is perceived as pitch), and *volume* (which is perceived as loudness) (ibid: 198). Usually, proximity signifies spatial proximity, as when we say that two sounds are near each other in the spatial regions of frequency, time, and loudness. Bregman (ibidem) states that we could also use a more neutral term, *similarity*. This term might be employed for the cases in which two sounds *sound* similar, but we are nonetheless unable to describe what the similarity consists of. In this paper, we will be using the two terms interchangeably.

The *common fate* principle establishes that, even though certain characteristics of the stream change over time, we are still able to perceive the stream as a persistent entity. The principle of *context* says that the way in which we experience the components of an auditory scene depends upon the specific role that these components play in the larger organization to which they belong. The dependency of the single elements to the context can also be captured by the principle of *exclusive allocation* or *belongingness*, according to which, even though there might be significant oscillations, a single element cannot belong to different auditory streams simultaneously and thus necessarily has to be attributed to a specific auditory stream. For example, there are cases in which a loud sound masks a low sound, preventing us from hearing it. We might still be able to detect the low sound when it is continuous and it is heard before the start of the loud sound and after it ends. In some particular circumstances, moreover, even if the softer sound is deleted during the brief loud sound, it is still heard as lasting during the interruption.

This is explained by the *closure* principle (ibid: 27). If one is surrounded by voices and overlapping noises, making the sonic environment especially chaotic, it can help to follow the voice of a friend. In this circumstance, the auditory system maintains the friend's voice in the foreground and the other sounds in the background, since, according to the principle of *organization*, it has the tendency to avoid disorganized experiences. Indeed, it is possible to opt for an alternative organization, namely, to move later from the voice of our friend, which then recedes to the background, to the rest of the sounds, which will then be perceived in the foreground. As Bregman writes: “at no time, is the experience ever free of either one organization or the other” (ibid: 199).

The auditory system employs the above-mentioned Gestalt principles for sequential and simultaneous integrations; the “materials” on which they operate are frequency, loudness, tone rate, spatial localization, and timbre. Let us see now in greater detail how exactly these principles operate when working sequentially and simultaneously.

### 3.1 Sequential integration

The cues that determine how individual sensory elements are grouped into auditory streams sequentially are the temporal rate of the sequence of sounds, the specific frequency of their tones, their spatial localisation, and their volume. In order to see how the cues of frequency and speed influence the sequential integration by using the principle of proximity, Bregman (1990, p. 17) discusses the case of a sequence of sounds constituted by a high-pitched tone A and a low-pitched tone B that alternate with a certain speed. The two tones have frequencies that are very far from each other. We would expect to hear an auditory stream constituted by tone A that alternates with tone B. Instead, we experience two separated sequences, one constituted by a succession of As and the other of a succession of Bs. In this case, the factors that influence the segregation of the two streams are the speed of execution of the two tones and the distance between their frequencies. Moreover, when in front of conflicting perceptual organizations, the fact that we necessarily listen to either one perceptual organization (the one that alternates between A and B) or the other perceptual organization (the one made up by a series of Bs and a series of As) is determined not only by the proximity of the frequency of the tones, but also by the principle of organization that establishes the fact that, in order to avoid unorganized experiences, the auditory system makes us hear either one or the other tonal series.

Other two grouping factors that are important for the sequential integration are spatial localisation and volume. When it comes to spatial localisation, we tend to group sounds that come from the same direction and the same distance and separate those that seem to come from a different direction and distance, again, in accordance with the proximity principle. The direction and distance from which sound issues are also important when, in a noisy room, we need to follow a particular conversation. These cues play an extremely effective part in sequential integration (Shinn-Cunningham, 2005), whereas they seem to be less effective in simultaneous integration. The proximity principle is also at work in the studies that illustrate that there is evidence that we tend to group loud sounds with other loud sounds into one stream, and low sounds with low sounds into another stream (ibid: 126).

Bregman and Rudnický (1975) conducted an experiment to analyze the application of the context, exclusive allocation, and proximity principles in determining sequential integration and employing pitch (Bregman, 1990, p. 14). The listener had to order (in terms of pitch) two tones, A and B, occurring in sequence, and they had to say whether the order was high-low or low-high. When A and B occurred as a pair of tones, in isolation, the listener ordered them correctly. But when the two tones “F” of equal frequency (“F” stands for “flankers”) were added to the sequence, one before the pair and one after, it was harder for the listener to order A and B correctly. The conclusion was that the specific order of A and B was lost when the two tones were embedded in a larger sequence.

The researchers also checked to see how the perception of the element AB would be affected by assigning the Fs to a different perceptual stream. To do so, they introduced a further group of tones C (which stands for “captors”) to see what happened when varying their frequencies. When those frequencies were much lower than the frequency

of the Fs, the Fs grouped with AB tones and, once again, the order of A and B was difficult to establish for the listener. When the C tones were near in frequency to the F tones, instead, the listener grouped them into the stream CCCFFCC. In this second case, the order of A and B became audible once more. As we have already discussed, tones tend to group with those that are the closest to them in frequency, therefore proximity explains the “behavior” of the Cs; the perception of the order AB depended on the allocation of Fs, as when AB was embedded in the sequence of Fs or when it was isolated. Again, in the test involving more Cs tones, AB became audible again, since they were on their own auditory stream (as they were at the start of the experiment), because the Fs tones were distanced so as to fall into a regular rhythmic sequence with the C tones and were also close in frequency to the C tones. The allocation of Fs had been altered because of the principle of context and, thus, the perceived auditory forms were changed, too. Focusing on how we hear Fs helps us understand the functioning of exclusive allocation; focusing on the audibility of AB inserted in a broader context exemplifies the principle of context; and focusing on the Cs and how they are related in frequency to Fs is illustrative of the proximity principle.

The principle of closure explains the illusion of continuity, which is another phenomenon relevant for sequential integration. There are different versions of the illusion of continuity (Miller & Licklider, 1950; Warren et al., 1972; Dannenbring, 1976; Bregman & Dannenbring, 1977; see Warren, 1982 for a review). The effect consists in the fact that, when rapidly alternating a tone with a noise burst, instead of hearing a sequence that is made up of tones and noise bursts, the listener hears a continuous sequence of tones without it being constantly interrupted by noise bursts. The principle of closure tells us that there are some “strong” perceptual auditory forms, such as a steady sound, that we tend to complete, that we tend to close—just as in vision we tend to complete “strong” forms such as circles. Therefore, there are cases, as in the illusion of continuity, in which the strong form, the steady sound, tends to be completed or, even, *reconstructed* by the auditory system, above and beyond the presence of distracting elements.

### 3.2 Simultaneous integration

With respect to simultaneous integration, similar cues to the ones influencing sequential grouping take place. The basic question of the simultaneous grouping process is: “How do we know which acoustic components have arisen simultaneously from the same physical event?” (Bregman, 1990, p. 221). The principle of proximity ensures that we tend to group auditory elements that resemble each other with regard to frequency and volume. Grouping together elements that resemble each other in frequency when we need to assemble different sounds into a single percept means that we unify those sounds (partials) that are harmonics of the same fundamental, namely, those sounds that are multiple integers of the lowest frequency that functions as an “attractor” for other frequencies that stand in relation to it. When different sounds are all multiple integers of this basic frequency, we hear all these frequencies as fused into a single sound (in light of the *harmonicity principle*). Moreover, we group simultaneous sounds



that share a spatial location, that is, we tend to group them when they issue from the same position.

The common fate principle applies when the components of a stream change simultaneously in volume or frequency so that we still hear the stream as such. Moreover, proximity in time, the fact that sounds start and end at the same time, influences simultaneous integration as well.

#### 4 The musical scene analysis and auditory chimeras

Bregman (ibid: chapter V) maintains that, given that music is made of sounds, its perceptual organization is likely to be governed by the same grouping principles that govern the organization of environmental sounds. This suggests that the primitive grouping cues and Gestalt principles that influence the perceptual organization of the ordinary auditory scene analysis and of musical scene analysis are alike. What this means is that, in the first case, the elementary auditory elements of loudnesses and pitches are grouped and segregated into environmental sounds or streams of environmental sounds, whereas, in the second case, the same auditory elements are grouped and segregated into melodies and harmonies.

The fact that the auditory system employs a method for grouping and segregating auditory elements that works sequentially and simultaneously within a musical context can be seen by considering the example of musical notation. Notes are written on the staff on a horizontal and a vertical dimension, in which the former represents sequential integration while the latter represents simultaneous integration. In the horizontal dimension, time is represented, while in the vertical dimension, what is represented is pitch. As Bregman writes: “[...] because much of the pitch and timing information is translated into the two spatial dimensions of the score, many of the perceptual groupings that we see in the score correspond to groupings that we hear in the music” (ibid: p. 456). In this section, I expand on Bregman’s claim by describing how the Gestalt principles that apply with respect to auditory cues determining the sequential and simultaneous integration of ordinary sounds *also* apply with respect to auditory cues determining musical sounds that generate auditory chimeras.

The first and crucial similarity between the ordinary scene analysis and the musical scene analysis is that, just as the task of the auditory system in the face of auditory *multiplicity* is to segregate and group auditory elements so that the auditory streams thus generated correspond to the sources from which they supposedly come from, likewise, the auditory system’s task in the face of musical sounds’ *multiplicity* is to segregate and group auditory elements to form musical streams originating from specific musical sources, be they real or imaginary. When listening to a symphony, for example, the auditory system has to segregate the musical streams corresponding to the different melodic lines expressed by different instruments, in such a way that each stream corresponds to the instrument producing it. Nevertheless, the melody-streams we segregate do not necessarily need to be carried out only by a single instrument. Therefore, the whole melodic compound comes from different instruments, so that we can say that each melodic fragment it is composed of (or originates from) real sources;

even though the compound *itself* is not attributable to any of the sources, we cannot say that neither the compound nor its source is imaginary.

There are many cases in which an entire melodic line is divided into melodic fragments, which pass from one instrument to another and form a single melodic stream, as in numerous sonatas for more than one instruments or concertos for two, three, and four instruments, as in many of Vivaldi's compositions or in Mozart's concertos for more than one instrument (such as the concerto for harp and flute K 299, the symphony for violin and viola K 364, or the concerto for piano and violin K 315f). In these cases, the primitive grouping takes place with the aim of attributing a stream to its source, but this source does not necessarily need to be a single one, since the carrier of the entire melodic line (the melodic chimera) is made up of multiple sources coming in succession. Musical chimeras occur not only when a melody generated by the sequential integration process comes from a succession of instruments, but also when different sounds occurring at the same time are assembled by sequential integration in a single stream, a harmonic compound, which does not belong, strictly speaking, to any single environmental source but is instead an emergent, new compound. This emergent compound, as for the melodic compound, is actually constituted by harmonic fragments generated by different real musical sound sources, but the compound itself is not attributable to any of those sources. Nevertheless, even though the compound itself is not attributable to any of those sources, we cannot say that neither the compound nor its source is imaginary.

Let us now describe how the sequential and the simultaneous integration work within the musical scene analysis. This will allow us to see how the same factors and Gestalt principles operating within the ordinary auditory scene analysis are also operative in musical contexts generating auditory chimeras.

#### 4.1 Sequential integration

Let us recall that the factors that affect sequential integration by virtue of the principle of proximity are the frequency of the tones, their rate, the spatial localization, and their volume. Both rate and frequency separation influence the integrity of the streams; if the rate and frequency jump rapidly back and forth between different frequencies, they will not be perceived as grouped. Vice versa, notes that are closer in frequency to one another "stick" together.

The proximity principle has been discussed by Bregman (1990, p. 17) in his analysis of the ordinary auditory scene by considering the case of a sequence of high-pitched tones (As) and low-pitched tones (Bs) that alternate following a certain speed. We can verify an application of the proximity principle by looking at several compositions of Baroque music, examining how the way in which we often distinguish one melody from another, including when it is played by multiple instruments, depends on the rate of the execution of the notes and their pitch. This technique was employed explicitly by many composers in order to make a specific melodic line easily recognizable and to put potentially competing melodies in the background. In fact, as it happens

with environmental sounds, also within a musical sound context, when there are conflicting perceptual organizations we necessarily listen to one perceptual organization of tones, as suggested by the principle of organization.

In order to analyze the rate of the tones required to ensure segregation, Dowling (1973) examined twenty recordings of Baroque music and found that, when the median tone rate is 6.3 tones per second, it is very difficult to get a clear segregation. As for the claim that smaller steps in pitch contribute to perceiving group tones as belonging to the same stream, instead, the psychologist Ortmann (1926), as early as the 1920s, analyzed compositions by Franz Schubert, Robert Schumann, Johannes Brahms, and Richard Strauss that have many melodic lines, and found that the smallest intervals between the notes were the most numerous, showing that composers are well aware of how to write a sequence of notes that will be correctly segregated (melody-wise) by the listener.

This applies also to melodies conducted by different instruments. We hear the theme of the fugue opening Anton Webern's orchestration of the *Ricercar* from "The Musical Offering" by Johann Sebastian Bach as a unitary melody, even though it has been carried out by different wind instruments (flute, oboe, English horn, clarinet, bass clarinet, bassoon, French horn, trumpet, and trombone). This is explained by the fact that the sequence of sounds is heard as a unitary melody even though it is produced by different sources, since the sounds composing the melody have smaller steps in pitch.

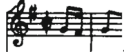
Just as, when talking about ordinary sounds, we tend to group sounds that come from the same direction and the same distance and separate those that seem to come from a different direction and distance, likewise, the direction and distance from which musical sound derives is equally important when segregating musical streams. When listening to an orchestra playing a symphony, we tend to segregate melodic lines that come from the same spatial region and have difficulties hearing sounds as constituting a unified perceptual event when they are constituted by sounds coming from different directions or at a different distance from the listener. When hearing melodic chimeras, too, we hear a melodic whole when the melodic fragments come from instruments placed not far from each other. The proximity principle applied to spatial location also explains why musicians playing different instruments make sure to sit on stage as close as possible, so that their spatial vicinity will contribute to the right perception of the melodic sequence.

As in the ordinary auditory analysis, it is equally true in the musical scene analysis that loudness does not segregate streams as effectively as frequency. Nonetheless, differences in loudness are still relevant, especially when they are sudden, since they demarcate a musical stream by signaling the beginning or the end of a louder event. This is the mechanism at work with respect to melodies carried by a single instrument as well as ones carried by multiple instruments: when the melodic fragments are uniform in loudness, they are grouped together to form a unified melody; when, instead, there is a sudden change in the loudness of one of the melodic fragments, then it is unlikely that this will be unified into a sole melody with other melodic fragments.

I mentioned the Bregman and Rudnický (1975) experiment because it is helpful in understanding the functioning of the principles of context, exclusive allocation, and proximity for sequential integration in the ordinary auditory case based on pitch variation (Bregman, 1990, p. 14). Let us recall that the experiment has a sequence of

two tones, A and B, which the listener is able to hear as a specific and isolated element and to order in terms of pitch when played in isolation. Nevertheless, when the element is anticipated and followed by the two tones “F” of the same frequency, the listener is no longer able to order A and B. This suggests that, when AB are included in a larger sequence, and if they are anticipated and followed by tones of equal frequency—which means that A and B are embedded in a broader perceptual context—we tend to focus on the entire sequence and lose AB specificity. Bregman and Rudnický also examined what happens when the Fs and AB are embedded into a broader perceptual stream. To do so, they introduced further tones, Cs, to create the larger stream CCCFABFCCC. When the Cs were much lower in pitch than Fs, Fs were grouped together with AB again, and the AB order was lost once more. When Cs were close in frequency to F, AB became audible again, and the listener tended to group the tones as forming the following stream CCF FCC.

The Gestalt principles of context, exclusive allocation, and proximity are employed in multi-instrumental music in which we have to be able to detect an element and to order the tones constituting it in terms of pitch. I will discuss a musical piece in which it is possible to individuate an element and to test how we hear it by virtue of the application of the principles of context, exclusive allocation, and proximity. The first movement of Bach’s 3rd Brandenburg concerto is constructed around a basic musical element (ab), which is composed of three notes with a difference of one semitone

between them, namely: . There are variations of this element, since the three notes might have a distance of a tone and can also be played by different instruments. This element, in order to be clearly heard, is often written in a way such that it is easily recognizable and so the listener can determine without any difficulty which note is higher. It almost never appears as embedded in a sequence within which it anticipates and is followed by notes of the same frequency, as in the FABF case. Indeed, Fig. 1 shows that the element ab is repeated twice for the violin and the viola section and once for the cello’s section. This fragment, made up of repetitions of this element, creates a melodic chimera that passes from the violin to the viola and, finally, to the cello section (AB). The entire sequence is not anticipated, and it is followed by single notes or a sequence of notes that have the same frequency. Moreover, inasmuch as the melodic chimera is also embedded within a broader perceptual sequence—as with the CCCFABFCCC case—we can still detect and order the tones composing the element, both when it is a single element as well as when it is forming a longer stream that is itself a melodic chimera made of different single elements. This is possible because the Cs elements do not prevent one from hearing AB. That is, since they are written in a way as to be closer in frequency to Fs instead of AB, AB is always somehow detectable.

We already saw how the principle of closure applies to environmental sounds through sequential grouping. There are many music compositions that take advantage of this principle in order to create continuity effects. For example, Deutsch (1999, p. 309) mentions that there are music examples in twentieth century guitar literature, such as Francisco Tárrega’s *Recuerdos de la Alhambra* and Augustín Barrios’ *Una Limosna por el Amor de Dios*, in which there are passages where, even though the same

**Fig. 1** The AB melodic chimera embedded in the CCFABFCC sequence from J. S. Bach's 3<sup>rd</sup> Brandenburg Concert. The element ab is repeated twice for the violin and the viola section and once for the cello section. This fragment made of repetitions of ab creates a melodic chimera that goes from the violin to the viola and, finally, the cello section (AB).

tone is quickly repeated many times and even regularly silenced and substituted with a different tone, the listener is still capable of perceptually generating the omitted tone. In other words, the listener “hears” these notes even when they are not being played, so that, in their experience, the strong form of a steady sound is maintained. This effect is also audible in all the musical samples in which the alternation between tones (heard as the steady sound with the distracting tones playing the role of noise bursts) are played by different instruments. For example, in the concertos for one mandolin and two mandolins by Vivaldi, the continuity effect is created by the interaction between the soloists and the orchestra. Or, in many transcriptions of Arcangelo Corelli’s *Follia* for two or more instruments, again, the continuity effect is quite recognizable and evident.

## 4.2 Simultaneous integration

The principle of proximity introduced for the sequential integration of ordinary sounds also applies to the simultaneous integration of musical sounds, which means that we tend to segregate two or more simultaneous musical sounds that resemble each other with regard to specific features, such as loudness and frequency. It applies to simultaneous sounds produced by a polyphonic instrument (such as pipe organ, piano or accordion) as well as to sounds coming from different sources. Just as the proximity principle applies in the case of the frequency of environmental sounds experienced by applying the harmonicity principle, the exact same principle is also operative when grouping simultaneous pitches coming from different sources as constituting harmonic chimeras. Most musical compositions are based on simultaneous sounds perceived as chords, namely as unified harmonic composites, as opposed to simultaneous sounds heard as disconnected, even when sounds are produced by different musical instruments.

To discuss only one of the many available examples of simultaneous sounds emitted by different musical instruments heard as a single chord, let us consider Mozart's quintet for piano, oboe, clarinet, horn, and bassoon in E flat [KV 452]. The first movement opens with all the instruments of the quintet playing at the same time and, instead of hearing a group of simultaneous sounds, we hear them as a unified chord: E-flat major. The harmonicity principle operates in this specific case as well, since the different sounds are all multiple integers of the basic frequency, which is E3 at 165 Hertz; this is why we hear all the frequencies as fused in a single sound.

We tend to group simultaneous ordinary sounds that share a spatial location, which happens when they come from the same position; likewise, we tend to group musical sounds that come from the same spatial region. This is true when the sounds are produced by a single polyphonic instrument, but also when they are produced by different instruments. This is the same principle at work in sequential grouping, and it explains why musicians, knowing that the sounds they play are supposed to fuse into a sole chord, tend to sit as close as possible, so that it is easier for them to create that unified sound.

The common fate principle applies when the components of a sound stream change in loudness or frequency simultaneously, so that we still hear the stream they belong to as such. Harmonic chimeras are determined by the same principle: when different musical instruments playing a chord change simultaneously in frequency and in loudness, we still hear the chord as produced by the same instruments we heard before the change.

Proximity in time, namely, the fact that sounds start and end at the same time, is another factor that influences simultaneous integration. Just as for ordinary sounds, in the case of musical sounds of different sources, synchronicity works as a temporal cue for sequential integration. When musicians start playing a chord and then stop playing all together, this strongly affects us in hearing the simultaneous sounds emitted by different instruments as a unified chord.

## 5 Some challenges to the perceivability of chimericity

In this section, I will discuss some worries that could challenge the rich view of audition I defend in this paper. In particular, against the perceivability of chimericity, one may argue that (1) following Reiland (2014), chimericity is a part of quasi-sensory/quasi-cognitive phenomenology, or (2) following Lyons (2005), that chimericity might consist in having a perceptual belief. These ways of representing properties may partially rely on perceptual mechanisms; therefore, it is not obvious that the involvement of grouping processes automatically allows to reject options (1) and (2). Another worry could be that, (3) since chimericity does not seem to belong to the same category as typical types of high-level properties related to general categories, emotions, and causality, further justification is required to explain why chimericity is a high-level property. Finally, (4) one might wonder whether the property of chimericity counts as sufficiently “rich” to challenge the sparse view. Let me start by addressing the first worry.

We do not necessarily need to be able to experience that the melodies and harmonies we detect are generated by different instruments. That is, we do not need to be *aware* of hearing auditory chimeras; we can also have the illusion that it is just one single instrument that generates them. What is essential for the experience of musical chimeras is that, thanks to a perceptual mechanism which I have tried to describe at length, we track a melody or a harmony by virtue of experiencing a feeling of unity, even though it is generated by different sound sources. We can exclude, thus, that chimericity is experienced as a quasi-sensory/quasi-cognitive phenomenology, and that it consists in the experience of having a perceptual belief.

Let me explain. Reiland (2014), following Brogaard’s *phenomenal* use of perceptual verbs (2013), introduces a distinction between perceptual events and *seeming* events. A seeming event might occur when we say, for example, that *x* is the event of something seeming to us that it is a pine tree. Seeming events are *sui generis* phenomenal events (ibid: 9) that are passive and conceptual while also representing objects as having certain properties. They share the quality passivity with perceptual events, as they just occur, which means that they do not cease to exist in the presence of a defeater. Conceptuality and the fact that seeming events represent objects as having certain properties is shared, instead, with judgmental states, as they are “made as a result of deliberation and involve making up our mind” (ibidem). As for the conceptual capacity, they represent an object as belonging to a conceptual category. Reiland takes seeming events as interfaces between perception and cognition, and he characterizes them as having a quasi-sensory/quasi-cognitive phenomenology. As I have already mentioned, in the case of chimeras we do not necessarily need to be able to experience that the melodies and harmonies we detect are generated by different instruments; therefore, we do not need to be aware of hearing auditory chimeras, which implies that we do not need to be able to employ concepts and represent objects as having certain properties. We can have the illusion that it is just one single instrument that generates chimeras, since what is essential for the experience of musical chimeras is that we track a melody or a harmony by virtue of experiencing a feeling of unity. Therefore, the experience of chimericity is not a seeming event: it is not equipped with conceptuality and the

capacity to represent an object as having certain properties. Consequently, it is not experienced as having a quasi-sensory/quasi-cognitive phenomenology.<sup>5</sup>

Let me turn now to the second worry, which is the issue of whether experiencing chimericity consists in having a perceptual belief. According to Lyons (2005), perceptual beliefs are those beliefs that are based on how things look, sound, taste, smell, or feel to us. For example, for Lyons, I could form the belief that it's raining, but this would be a perceptual belief only if it resulted from my looking out of the window. It would not be a perceptual belief if it resulted from my listening to the weather report on the radio (ibid: 250). Or, imagine a chicken sex expert and a novice placed in front of a male chicken: whereas the latter forms the perceptual belief that he is looking at a male chicken by looking at the chicken, the novice is simply guessing that the chicken is male; therefore, she cannot have a *perceptual* belief. Chimericity is a perceivable property that does not necessarily generate the correspondent perceptual belief because we do not need to be aware of hearing auditory chimeras; we can also have the illusion that it is just one single instrument that is generating what we are hearing. Given that what is essential for the experience of musical chimeras to occur is that we track a melody or a harmony by virtue of experiencing a feeling of unity, we perceptually experience chimeras ipso facto by experiencing a feeling of unity. Then, we could eventually form the perceptual belief that we are hearing a chimera, but this is a further step which should not necessarily be taken. Therefore, I can exclude that experiencing chimericity consists in having a perceptual belief.

Concerning the third worry, according to which further justification is required to explain why chimericity is a high-level property, given that it does not seem to belong to the typical types of high-level properties related to general categories, emotions, and causality, Nanay (2012) argues that properties such as being edible or being climbable are not reducible to more basic properties as they have been evolutionarily useful for our ancestors to perform actions. The case of chimericity is different from the properties of being edible or being climbable since it does not seem to have had such a strong evolutionary role. Nevertheless, even though chimericity does not seem to have had this role, being able to perceive the category of musical chimeras is crucial for the appreciation of music, since it allows the listener to grasp some of the essential musical features of a vast majority of musical pieces, namely, the melodies and the harmonies that constitute the basic elements responsible for musical expressivity.

Given the disanalogy between action properties and chimericity, though, and in order to argue that chimericity is a high-level property not reducible to lower-level properties, I follow Skrzypulec's strategy (2019). He argues that recognition-based properties (R-properties) are irreducible to low-level properties by discussing how they are not reducible to shape Gestalten. My point is to show that chimericity is not

---

<sup>5</sup> I have claimed that chimericity is not experienced as a "quasi-sensory/quasi-cognitive" phenomenology. Nevertheless, one could further argue that it is the feeling of unity that has a "quasi-sensory/quasi-cognitive" phenomenology and not chimericity per se. Now, if the feeling of unity reveals the presence of chimericity because its content is constituted by the property of chimericity—which, as I have argued in this paper, is a perceivable property—then the feeling of unity is ipso facto experienced as perceivable as well. This cannot be experienced, indeed, as a "quasi-sensory/quasi-cognitive" state; otherwise, this would be in contradiction with the fact that its content is a perceivable property.



reducible to the low-level properties of loudness, timbre, and pitch. In order to do that, I will consider each low-level property in turn.

Let me start now by discussing whether chimericity can be reduced to loudness, that is, whether variations in terms of loudness determine different chimericity properties. Imagine listening to melodic fragments coming from different musical instruments and assembling them as forming a melodic chimera. Variations in the loudness of sounds are usually due to changes in the location of the musicians or to the subjective choice of the musicians to change the volume of their playing. Nevertheless, variations of this kind can hardly divert one from hearing a melodic chimera, as they would not impact the sense of unity that characterizes it. The same happens with harmonic chimeras: if one or more sounds composing them changes in loudness, that would not impact the perception of harmonic chimeras to the point of preventing us from hearing them as a unitary whole.

What about timbre? Can we say that changes in timbre determine different chimericity properties? Given that each melodic fragment (or harmonic sound) is coming from a different instrument, there is always a timbral variation. Nevertheless, in the cases in which the timbres of the instruments are very close to each other—as in the case of the timbre of the violin and the viola—this variation does not prevent from hearing the melodic (and harmonic) chimera. In the case in which instruments differ from each other in terms of timbre—as in the case of a piano and a violin,—we obviously hear a subtle change of timbre, which confers a specific, nuanced allure to the melody without distracting us from hearing the unitary melodic (or harmonic) compound.

Let us turn to pitch. In order to hear a specific sequence of sounds as a melody and a specific group of simultaneous sounds as a harmonic compound, we need to take into account the pitch of these sounds and preserve the distance between them. Nevertheless, in order to hear a melodic or harmonic chimera we do not need to grasp a sense of unity, which remains unaltered despite changes in pitch and variations in the distance between notes. For example, a melodic chimera can be perceived as such despite changes of octaves, chromatic variations, insertions of grace notes, trills, and changes of major or minor third between two notes. The same happens for harmonic chimeras: not only changes in octaves, but also many different changes in terms of the distances between the notes grouped into the compound (i.e. changing a minor second into a major second, a minor third into a major third, or a minor sixth into a major sixth) would not affect the sense of unity characterizing the chimera.

Finally, the last challenge to my view is that one might wonder whether the property of chimericity counts as sufficiently “rich” to question the sparse view. Yes, it does, as I think that when discussing audition, we should focus on a specific property (i.e. being a gendered voice, being a musical sound, being an environmental sound, being a beautiful sound, being the sound of a language we understand, and so on), and investigate the experience of that property. If this experience has the characteristics that are usually attributed to perceptual experience, and if that experience is not reducible to the experience of more basic or low-level properties, then we can tell whether our auditory experience of the property is rich. Therefore, we can evaluate whether our auditory experience is rich only with respect to a specific property. It is plausible that audition is thin with regard to the property of being a gendered voice (as our experience of this property does not have the characteristics usually attributed to

perceptual experience, and can be reduced to the experience of the low-level properties of pitch, loudness, and timbre), but rich with respect to the property of being the sound of a language we understand (as this experience has a genuinely perceptual character, which cannot be reduced to lower-level properties). The same is true of vision: the content of visual perception might be thin with regard to specific properties (i.e. being a beautiful human face or a stunning landscape) and rich with regard to other properties (i.e. being a natural or artificial kind, being an edible object, being a specific emotion expressed in human faces).

## 6 Conclusion

I have shown that the grouping property of chimericity is a property that we can auditorily perceive, and I have addressed the most prominent worries that might have challenged this idea. My argumentative strategy is based on the fact that, if the properties of being environmental sounds are perceivable by virtue of a perceptual mechanism (i.e. the primitive grouping operating via Gestalt principles), then we have good reasons to think that the properties emerging as the result of the same perceptual mechanism employing similar Gestalt principles should be perceivable as well. Given that chimericity emerges as the result of the primitive grouping employing Gestalt principles, we can conclude that it is indeed a perceivable property.

I did not delve into the issue of *which* Gestalt principle is more effective for the segregation and grouping of sounds in either the ordinary scene analysis or the musical scene analysis. Therefore, I cannot rule out that, even though the principles of the primitive grouping acting in both analyses are the same, they might weigh differently when parsing the ordinary and the musical auditory scenes. Nevertheless, even though future research could show that such principles weigh differently in the two contexts, it will not undermine my aim to show that we can perceive chimericity, since they will be principles with different roles but that are both always at work at the perceptual level.

Moreover, I did not aim to show that musical experience and the experience of environmental sounds are exactly the same perceptual experience, since showing that a musical property is perceivable, which is my aim here, does not mean to show that musical experience and the experience of environmental sound share the same phenomenology. That is a question for a future occasion.

**Acknowledgements** Earlier versions of this paper have benefited from discussions with the audience attending conferences in Austin, Jerusalem, L'Aquila, Oslo, and Strasbourg. I am particularly grateful to Stefano Ercolino, Karma Ben-Johanan, Krisztina Orbán, Susanna Siegel, Assaf Tamari, Preston Werner, Hong Yu Wong, and the anonymous referees for their insightful comments.

**Funding** Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative

Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research*, 89(1), 560–572.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. The MIT Press.
- Bregman, A. S., & Dannenbring, G. L. (1977). Auditory continuity and amplitude edges. *Canadian Journal of Psychology*, 31, 151–159.
- Brogaard, B. (2013). Do we perceive natural kind properties? *Philosophical Studies*, 162(1), 35–42.
- Bregman, A. S., & Rudnicki, A. (1975). Auditory segregation: Stream or streams?. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 263–267.
- Butterfill, S. (2009). Seeing causes and hearing gestures. In K. Hawley & F. Macpherson (Eds.), *The admissible contents of experience* (Vol. 201, pp. 36–59). Wiley.
- Casati, R., & Dokic, J. (1994). *La philosophie du son*. Nîmes.
- Casati, R., Di Bona, E., & Dokic, J. (2013). The Ockhamization of the event sources of sounds. *Analysis*, 73(3), 462–466.
- Clark, A. (2000). *A theory of sentience*. Oxford University Press.
- Dannenbring, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology*, 30, 99–114.
- Deutsch, D. (1999). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music* (pp. 299–348). Academic Press.
- Di Bona, E. (2017). Towards a rich view of auditory experience. *Philosophical Studies*, 174(11), 2629–2643.
- Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, 14, 37–40.
- Dretske, F. (1995). *Naturalizing the mind*. The MIT Press.
- Fish, W. (2013). High-level properties and visual experience. *Philosophical Studies*, 162, 43–55.
- Helton, G. (2016). Recent issues in high-level perception. *Philosophy Compass*, 12, 851–862.
- Kulvicki, J. (2008). The nature of noise. *Philosophers Imprint* 1–16.
- Lyons, J. (2005). Perceptual belief and nonexperiential looks. *Philosophical Perspectives*, 19, 237–256.
- Masrour, F. (2011). Is perceptual phenomenology thin? *Philosophy and Phenomenological Research*, 83(2), 366–397.
- Miller, G. A., & Licklider, J. C. R. (1950). Intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 22, 167–173.
- Nanay, B. (2011). Do we see apples as edible? *Pacific Philosophical Quarterly*, 92(3), 305–322.
- Nanay, B. (2012). Perceptual phenomenology. *Philosophical Perspectives*, 26(1), 235–246.
- Nudds, M. (2009). Sounds and space. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essay* (pp. 69–96). Oxford University Press.
- O'Callaghan, C. (2007). *Sounds: A philosophical theory*. Oxford University Press.
- O'Callaghan, C. (2009). Sounds and events. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 26–49). Oxford University Press.
- Ortman, O. (1926). On the melodic relativity of tones. *Psychological Monographs*, 35(162), 1–47.
- O'Shaughnessy, B. (2000). *Consciousness and the world*. Oxford University Press.
- Pasnau, R. (1999). What is sound? *Philosophical Quarterly*, 49, 309–324.
- Pasnau, R. (2009). The event of color. *Philosophical Studies*, 142, 353–369.
- Peacocke, C. (1992). *A study of concepts*. MIT Press.
- Perkins, M. (1983). *Sensing the world*. Indianapolis.
- Reiland, I. (2014). On experiencing high-level properties. *American Philosophical Quarterly*, 51(3), 177–187.

- Shinn-Cunningham, B. G., (2005). Influences of spatial cues on grouping and understanding sound. In *Proceedings of the Forum Acusticum* (pp. 1539–1544).
- Siegel, S. (2009). The visual experience of causation. In Hawley & Macpherson (Eds.) (pp. 150–171).
- Siegel, S. (2010). *The content of visual experience*. Oxford University Press.
- Skrzypulec, B. (2019). Perceptual kinds as supervening sortals. *Pacific Philosophical Quarterly*, 100(1), 174–201.
- Sorensen, R. (2009). Hearing silence: the perception and introspection of absences. In M. Nudds & C. O’Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 126–145). Oxford University Press.
- Strawson, G. (1994/2010). *Mental reality*. The MIT Press.
- Tye, M. (1995). *Ten problems of consciousness*. The MIT Press.
- Warren, R. M. (1982). *Auditory perception: A new synthesis*. Pergamon.
- Warren, R. M., Obusek, C. J., & Ackroff, J. M. (1972). Auditory induction: Perceptual synthesis of absent Sounds. *Science*, 176, 1149–1151.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.