arts

MDPI

Article

Multivariantism of Auditory Perceptions as a Significant Element of the Auditory Scene Analysis Concept

Adam Rosiński

Faculty of Arts, University of Warmia and Mazury in Olsztyn, 10-007 Olsztyn, Poland; adam.rosinski@uwm.edu.pl

Abstract: The concept of auditory scene analysis, popularized in scientific experiments by A. S. Bregman, the primary architect of the perceptual streaming theory, and his research team, along with more recent analyses by subsequent researchers, highlights a specific scientific gap that has not been thoroughly explored in previous studies. This article seeks to expand on this concept by introducing the author's observation of the multivariant nature of auditory perception. This notion suggests that listeners focusing on different components of an auditory image (such as a musical piece) may perceive the same sounds but interpret them as distinct sound structures. Notably, even the same listener may perceive various structures (different mental figures) when re-listening to the same piece, depending on which musical elements they focus on. The thesis of multivariantism was examined and confirmed through the analysis of selected classical music pieces, providing concrete evidence of different interpretations of the same sound stimuli. To enhance clarity and understanding, the introduction to multivariantism was supplemented with graphic examples from the visual arts, which were then related to musical art through score excerpts from the works of composers such as C. Saint-Saëns, F. Liszt, and F. Mendelssohn Bartholdy.

Keywords: perception; listening; auditory scene analysis; music theory; psychoacoustics; musical acoustics

Citation: Rosiński, Adam. 2024. Multivariantism of Auditory Perceptions as a Significant Element of the Auditory Scene Analysis Concept. *Arts* 13: 180. https://doi.org/10.3390/arts13060180

Academic Editor: Ivana Medić

Received: 20 August 2024
Revised: 20 November 2024
Accepted: 5 December 2024
Published: 9 December 2024

1. Basics of Perceptual Streaming Theory

Auditory scene analysis is a fundamental process occurring in the perception of music, comprising the segregation and integration of sound components with the purpose of forming a cohesive perceptual experience in a given environment based on factors like timbre, rhythm, and register (McAdams 2024; Wang and Chang 2008). Seeking similarities (integration) or disparities (segregation) in the instrument timbres, melodic lines, harmony, and rhythm is an elementary process of reasoning about a musical piece (Barrett et al. 2021). This process is very complicated and is crucial when trying to understand complex musical compositions, as it enables listeners to focus on specific auditory streams and objects while ignoring the rest, which shows the intricate nature of the auditory scene analysis in the perception of music (Trainor 2021; Kondo and Kashino 2009; Pressnitzer et al. 2011).

The formation of perceptual images is influenced by factors such as the clarity of the melody and the processing of harmonic information, which can activate specific areas of the brain (Spada et al. 2014). The perception of music involves several cognitive functions, including acoustic analysis, auditory memory, and auditory scene analysis, which are essential for processing musical syntax and semantics (Koelsch 2011), while the perception of musical rhythms relies on the ability to make time predictions, which is a basic aspect of time processing, significant in auditory scene analysis (Rajendran et al. 2018; Hove et al. 2014). Musical spectral and temporal structures memorized by the listener serve the purpose of gathering data about sounding objects and their location (McAdams 2024; Trainor 2015). Perceptual streams created in this way during a live concert or when listening to a recording leave a certain interpretive mark on the perceptor's mind. Depending on the way in which the listener receives the sound structures, it is possible to capture various shapes and figures in his mind. The greatest factors influencing the formation of different shapes and figures are:

- The perceptor's attention and focus.
- Fragments on which the listener is focused in a given moment.
- Dividing heard sounds into two or more layers connected, for example, with the main melody and harmony (background sounds).
- The occurrence of masking sounds may distract the listener's attention, which might cause the main figure to be transformed into the background or vice versa (Huron 2016).
- A shift of attention into different perceptual elements is noticed in music, which, by some degree of similarity (several common features) with other sounds, can cause "switching the listener's focus" and the formation of totally new perceptual shapes and figures.

Polyphonic music is an interesting example of auditory scene analysis that is very significant in the interpretation of incoming sounds. Carefully listening to individual voices and melodies sensorily engages the listener in the context of many sounds, which can be led together, separately, convergently, or in a certain correlation, which may cause the formation of diverse mental representations in the perceptor's mind (Wright and Bregman 1987; Huron 1991; Janata et al. 2002; Disbergen et al. 2018; Siedenburg et al. 2021). However, not only polyphonic works can be used to illustrate the principles of auditory stage analysis. One such example is *Boléro* by M. J. Ravel, in which the mutual adoption of the melody and accompaniment by numerous musical instruments results in the listeners receiving various mental representations (Kashino 2006; McAdams et al. 2004).

Considering studies related to perceptual stream analysis, this problem can be extended to the realm of music, with the aim of identifying moments significant for both science and music and comparing theoretical assumptions with specific examples from musical literature. The analysis presented here seeks to find evidence of perceptual streams in selected musical excerpts, rather than solely in test sound sequences created in artificial laboratory environments.

2. Multivariantism of Auditory Perceptions

The stimuli reaching the healthy perceiver are always being processed and interpreted by the observer's brain, regardless of the type of incoming information, whether it is visual, gustatory (olfactory), tactile, or auditory. The described processing occurs in the central nervous system and is based on receiving signals incoming from the environment, memorizing them, transforming them, and eventually reintroducing them into the environment in the form of behavior. Cognitive processes include, among others, attention, awareness, perception, memory, thinking, and reasoning. The cognitive construct created as a result of these processes is used to build a mental image of the received stimuli. The cognitive processes depend on various factors related to the past experiences of the person receiving the sensory stimuli. Musical education is one of the elements that significantly influences the processing and organizing of received stimuli into perceptual streams. Thus, different individuals may be characterized by different sensitivity levels during the stimulation of their hearing organs with the same sound material (Rosiński 2021, 2023; Hubbart 2010; Pecenka and Keller 2009; Carey et al. 2015).

In this analysis, the occurrence of multivariantism is treated as a unique ability of listeners that can be observed in certain selected perceptual cases. This ability enables the perceiver to recognize different mental figures despite the fact that the hearing organ is constantly stimulated by, for example, the same excerpt from the musical piece. The number of variants that the listener is able to observe depends on the complexity of the work and the cognitive strategies used by the listener when analyzing the sounds incoming at a given moment. Multivariantism is a perceptual phenomenon showing that a different way of perceiving the same sounds is not a listener's mistake or a mental illusion but rather a certain potential of perceivers that should be further explored and understood in order to be used consciously.

3. Theoretical Foundations for Understanding the Occurrence of Multivariantism in Music

Humans are visual learners (Russ and Neal 2016; Poeppel and Overath 2012; Scheffer and Manke 2018), as more than 50% or even over 70% (Kirwan 2016; Cf. Pearson 2019) (according to different measurements by different researchers) of the cerebral cortex is responsible for processing just the visual part of incoming information (Dykes 2020). Such a large part of the brain responsible for the reception of visual stimuli is a result of human evolutionary development (Keller 2016; Devinsky and D'esposito 2004; Thompson 2011). For this reason, considering the importance of vision and its dominance over other senses, considerations referring to the multivariantism of auditory perceptions will be introduced using visual examples. In this case, it will be much easier for the reader to understand the formation of diverse mental shapes and figures that may appear during the analysis of music, as the presented graphic examples can be treated universally.

Figure 1a presents a field composed of eight right isosceles triangles. Four triangles are light in color, and the other four are dark. As a result of focusing on particular parts of the image, the perceptor may obtain two different perceptual variants of the same image. Figure 1b presents a "windmill" with dark blades, spinning in one direction, while Figure 1c shows a "windmill" with light blades, spinning in a direction contrary to the one in Figure 1b. In this case, it can be observed that a single graphic image may be the source of two different mental images (perceptual variants).

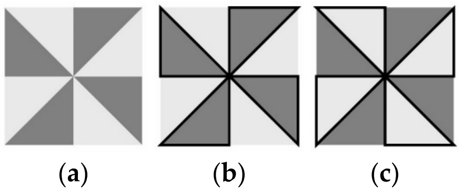

Figure 1. The field presented in (**a**) can be interpreted by the receiver in two different ways, as shown in (**b**,**c**). Source: own research.

The described illustrations can be compared to a very simple musical piece comprising melody and harmony. In this case, the melody and the harmony should both be led alike, in the form of overlapping sounds of an instrument/voice, or instruments/voices—at the same time or of similar, possibly partially the same pitches. If the sounds are led in this way in a musical piece, they enable the listener to perceptually exchange the main figure with the background, thanks to which, sometimes the background (harmony) may also become the main figure (melody), while the main figure (melody) can be interpreted as the background (harmony).

Figure 2a presents another field composed of eight right isosceles triangles. Four triangles are light and the other four are dark. Different positions of the triangles in relation to each other can make the perceptor observe totally different figures, depending on which elements of the presented image he will focus on. Different arrangements of the inner figures (triangles) make it possible for the perceptor to observe even six variants of the same image. Figure 2b,c presents the ways that the perceptor may spot a square on the image. In Figure 2b, the perceived square "lays" on the side, while in Figure 2c, a smaller square "stands" on the corner. Figure 2d,e shows other possible figures that may be observed in the image: a light hourglass shape (Figure 2d) and a dark variant of the same shape (Figure 2e). Figure 2f,g presents more interpretations of the same image—two double-sided "arrows" with different orientations, the first one composed mostly of dark triangles and the second one of light triangles.

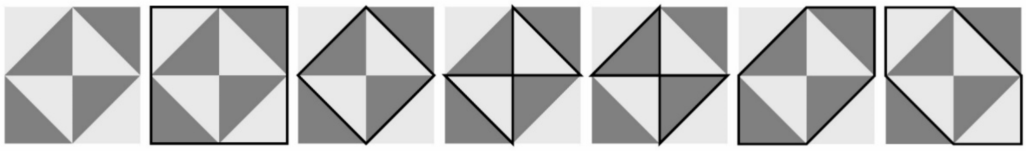

Figure 2. The field presented in (**a**) can be interpreted by the receiver in a few different ways, as shown in (**b**–**g**). Source: own research.

The presented illustrations can be compared to more advanced music, with a larger number of musical instruments or voices present at the same time. In this case, there is no simple change in the background or the main figure based on similarities. Common or differentiating features observed by the listener in the voice ranges, the instrument timbres, the spatial location, the harmony, etc., become the source of multiple perceptual variants of new, previously unseen shapes.

Figure 3a shows a widened field in comparison to Figures 1a and 2a. A larger number of triangles build this field, which helps obtain many more perceptual variants. The variants presented in Figure 3b–ad can be compared to the case of very complex music, where multiple intertwining and overlapping voices and musical instruments make it possible to experience numerous diverse perceptual images. Not every listener is able to spot all the perceptual variants, as it depends on one's attention (Bregman 1990), attitude (van Noorden 1975), and musical education (Rosiński 2021), all of which influence the process of grouping sounds into perceptual streams.

Figure 3b,c presents dark "windmills" on a light background. Depending on the elements of the image that are focused upon, one may observe "windmills" of the same color but "spinning" in opposite directions. An analogous situation can be observed in Figure 3e,f. In this case, the perceptor sees light "windmills" on a dark background. In Figure 3d,g, a dashed line indicates a shared element, which may diversify the observed perceptual image and split it into two different subvariants of the "windmills". In this case, the perceived "windmills" are of the same color, but they "spin" in opposite directions. Figure 3h,i divides the main field into four or sixteen pieces—depending on what perceptual element the receiver bases his/her observations on. Figure 3j,k shows interpretations of the image selecting four squares arranged in a way forming a diagonal of a whole field. Figure 3l presents the interpretation of the image as the "X" symbol. Figure 3m shows a very similar observation. However, here, the viewer spots the four squares in the corners, at the same time ignoring the square positioned centrally (marked with an "X" symbol). Figure 3n shows the case in which the perceptor also sees the square in the middle (compared to Figure 3m, where it did not happen). The interpretation of Figure 3o,p makes it clear that perceptual images in the form of diagonal lines, similar to half of the "X" symbol, can also appear. Figure 3q,r presents two double-sided "arrows" oriented in two different directions, very similar to the shapes shown in Figure 3o,p. In both Figure 3s,t, there are four "hourglasses"—in this case, the perceived shapes vary in color and orientation. Figure 3u,v shows "hourglass" shapes arranged on a diagonal of a whole field. This time, on each image, two of the three "hourglasses" are of the same color and one is different. Figure 3w,x presents how, when analyzing the image, the viewer may spot many triangles in two colors laid on diagonals of a whole field. These two figures could also be interpreted not as triangles but, for example, as small squares with a diagonal line—this or another interpretation is a form of an existing subvariant. Figure 3y shows that the entire field can be divided in the perceptor's mind into a series of thirty-two light and dark triangles. Figure 3z–ad presents triangles with larger areas than the ones observed before. In each interpretation, the position of the described triangle is different. Furthermore, three additional sub-variants have been marked with the letters "A", "B", and "C", as they may arise in the receiver's mind when interpreting different parts of the image (Brewer and Loschky 2005; Kornmeier and Bach 2012; McKenna 2020; Macpherson 2006).

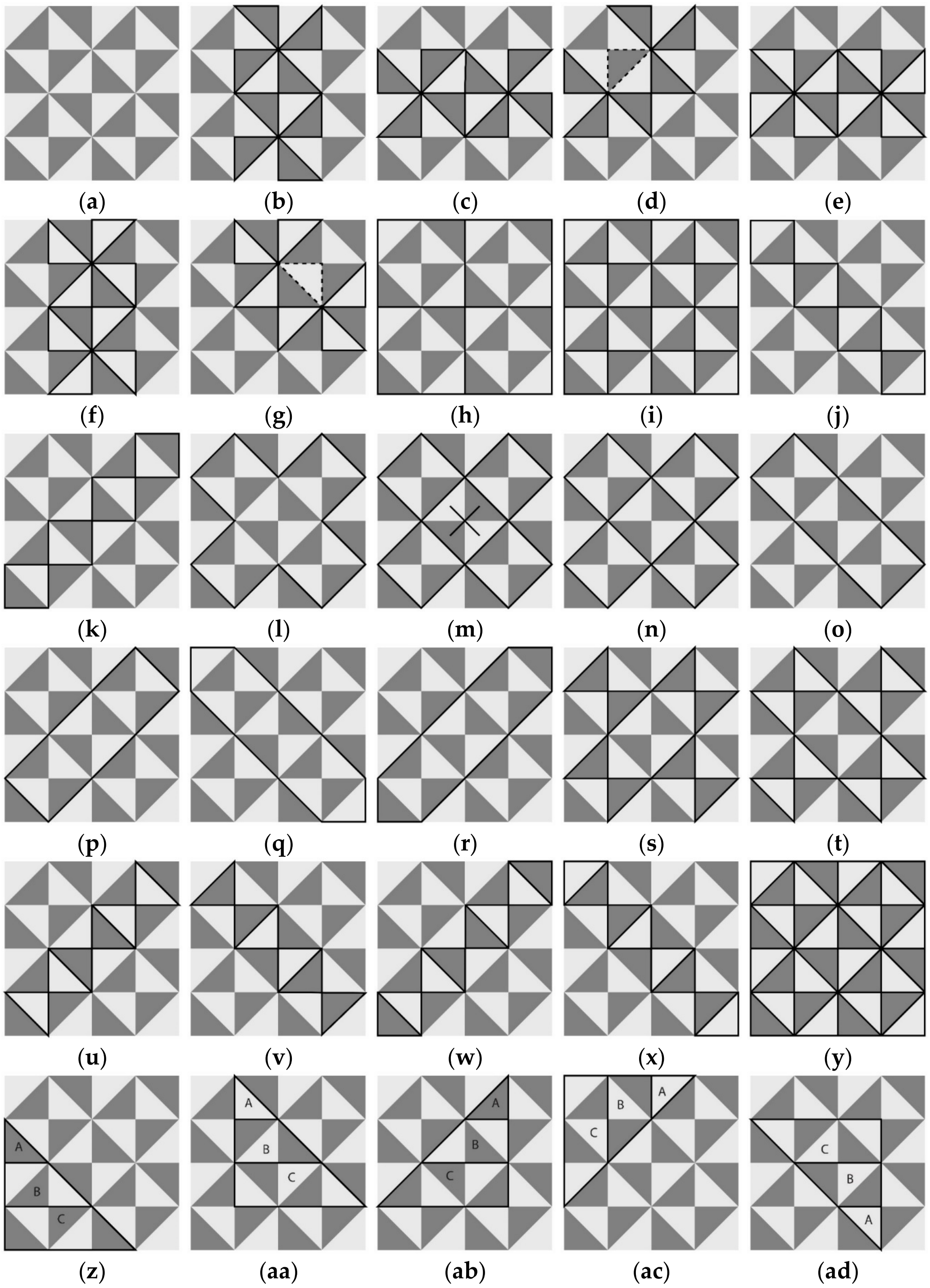

Figure 3. The extended field presented in (**a**) can be interpreted as very many perceptual variants (**a**–**ad**), depending on the element that the viewer focuses his attention on in a given moment. Source: own research.

It is worth mentioning that the author did not demonstrate all the variants that could arise as a result of the analysis. His primary goal was to prove and make it clear that a simple graphic pattern, such as the one presented, can be a cause and, at the same time, the main source of the formation of diverse interpretations, which is evidence of the occurrence of multivariantism of perceptions.

Figures 4–10 show how diverse perceptual variants can be observed by a receiver whose hearing is stimulated by the same melody. In order to present the reception of various forming perceptual figures in the form of perceptual streams in the listener's mind, an example was used—an excerpt from measures 3–4 of Johann Sebastian Bach's *Toccata and fugue in d minor (Dorian)*, BWV 538, *Toccata*.

Figure 4 shows variant No. 1. All the sounds are integrated by the perceptor into one undivided perceptual stream. This way of sound reception is characteristic of people with no musical education, who are not able to analytically perceive and mentally arrange individual pitches from the maze of many sounds coming to them (Rosiński 2021, 2024).

Figure 4. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for variant No. 1. Source: own research.

Figure 5 presents a situation in which the perceptor hearing sounds integrates them in the form of variant No. 2. In this case, depending on the listener's focus, attention, and the way in which the piece is performed (interpretation), the analyzing person can perceive every other note from the sixteenth-note passage as one perceptual stream. The highest sounds selected from the passage may be a basis for the mental formation of the melody (the first stream), while all the other sounds can be assigned to the second stream, constituting the harmony of the musical work (Salamon et al. 2012; Palmer 2015).

Figure 5. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for the variant No. 2. Source: own research.

Figure 6 presents the third perceptual variant, which is a mirror image of variant No. 1. In this case, the repeated A3 notes in the sixteenth-note passage can be integrated as a second perceptual stream, while all the higher notes will make up the harmony of the piece, forming the first perceptual stream. The second and third variants together are a great example of swapping the main figure with the background depending on the reception scheme used by the perceptor during the listening of the musical material. The main melody can join the harmony, or one of the elements of the harmony (in this case, the lowest voice) can be mentally transformed and perceived in the listener's mind as the melody (Salamon et al. 2012).

Figure 6. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for variant No. 3. Source: own research.

Figure 7 shows another fourth perceptual variant. The sixteenth-note passage, comprising the higher notes (played by the organist with his right hand), forms the first integral perceptual stream, while the second stream is made of the lower notes (played with the left hand). This way of perceiving stimuli arises when the perceptor divides both passages into two similar halves, according to the musical notation and the way of performing it with the left and right hands. In the case of instrumentalists who are familiar with this piece and have presented it before, the performance issues responsible for the motor skills of the hands while playing may suggest such a way of segregating sounds into two perceptual streams (Pressnitzer et al. 2011).

Figure 7. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for variant No. 4. Source: own research.

Figure 8 presents a very interesting fifth perceptual variant—the notes that are lined up vertically in the score (sounds in harmony) form two separate perceptual streams. Perceived streams are intertwined and heard alternately; hence, the listener's attention must be strongly focused on individual sounds in order to achieve continuity of the perceptual stream. The first stream consists solely of the sounds with a smaller interval in between, in other words, being "inside" or "in the middle". The extreme—the lowest and the highest sounds, being "outside" of the composition—are perceived by the listener as the second perceptual stream (McAdams and Bregman 1979).

Figure 8. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for variant No. 5. Source: own research.

Figure 9 shows a very interesting acoustic phenomenon, which can be perceived as variant No. 6. The perceptor combines sounds into perceptual streams on a crossover basis. One lower sound, played in the sixteenth-note passage of the left-hand part, is combined into one perceptual stream with the next lower sound of a sixteenth-note passage (not from the same harmonic vertical), which was written in the right-hand part, etc. A second perceptual stream is then built based on the pattern described above; however, in this case, the stream is formed by cross-combined higher sounds. This type of sound integration into individual streams occurs only during high-tempo performances due to the limitations of human perception. The fast tempo makes the perceptor unable to accurately and with great precision grasp the pitch of the consecutive individual tones (Repp 2009; Botte et al. 1997). The mentioned tempo causes a kind of blurring of the focus of a given perceptual image (problem with recognizing the exact arrangement of sounds), which causes the listener to be able to combine sounds on the basis of other perceptual criteria that are not necessarily correct or well understood by the listener.

Figure 9. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for variant No. 6. Source: own research.

In Figure 10, the formation of four different perceptual streams can be observed. The seventh variant is very difficult to observe because the listener can focus only on two streams at the same time (Zawadzka-Gołosz 2018), but by modifying and switching his attention to other occurring sounds in a certain way, he is able to perceive the other occurring streams (Załazińska 2016; Strumiłło 2012). This is the way in which the sound material is perceived by professional musicians, who, during their specialized musical education, develop certain listening skills and qualities that allow them to perceive sounds analytically. In this case, the listener is able to perceive as many as four distinct streams, each of which can be interpreted as one of four separate voices (Bigand et al. 2000).

Figure 10. Johann Sebastian Bach, *Toccata and fugue in D minor* (*Dorian*), BWV 538, *Toccata*, a fragment of measures 3–4, mental representation for variant No. 7. Source: own research.

The author acknowledges that it is impossible to demonstrate all potential perceptual variants that may emerge for different listeners when exposed to the same auditory stimuli. This is due to individual characteristics, such as sensory sensitivity, musical education, fatigue, mindset, or the method of listening to the musical material, all of which can influence how stimuli are perceived. The primary objective is to demonstrate that even a simple sequence can be interpreted in diverse ways by listeners and to highlight the multivariant nature of perception. Recognizing and uncovering these aspects represents a critical step in fostering this skill, thereby contributing to a deeper understanding and further advancement of auditory scene analysis methodologies.

4. Equipment Used During the Listening Sessions

The listening station used for the analysis comprised the following elements:

- A portable computer: *HP Pavilion Aero 13-be0802nw, with Microsoft Windows 11* operating system installed.
- An audio interface working as a sound card: *Steinberg UR824*.
- Wired headphones: *Beyerdynamic DT 770Pro*, a version with an impedance of 80 ohms.

The volume of the musical pieces played was determined during the test listening session of the sound material—pilot tracks. The audio path gain was set to a value with which the sound volume provided maximum comfort when listening to various musical pieces (Rosiński 2021). The maximum loudness level was equal to 95 phons during the loudest parts of played music, while the average loudness was approximately 73 phons. The sample rate and resolution parameters set in the audio interface during sound playback were equal to 44.1 kHz and 16 bits.

5. Listening Method, Listening Sessions, and Audio Material

The inconsistency of the musicians' performances (interpretations) and the phonography made it necessary to select fragments of works from many different records in order to unify and bring the different recordings to one sonic level, not only technically (in the area of sound engineering), but also considering:

- The inconsistent acoustics of different concert halls.
- Different timbral characteristics of each individual instrument.
- The diversity in the performance and interpretation of various musical works (De Poli 2004).

The listening sessions were held in October 2023. Each session lasted for no more than two hours. A maximum of one listening session per day was planned so that attention and focus could be kept at a high level in order to reduce the number of potential perceptual errors resulting from the listener's fatigue (Simon et al. 2020; Jain and Nataraja 2019; Jain et al. 2022; Kin and Dobrucki 2016; Moore et al. 2017).

The author of this publication personally conducted the auditory analysis of all the referenced recordings. Drawing on prior theoretical and auditory expertise acquired through earlier scientific experiments on the perception of auditory streams, the author examined various interpretative possibilities of musical works as perceptual streams that listeners might experience. Similarly to the approach adopted by music theorists—who focus on key elements of a musical work, such as notated melodies, harmonies, rhythms, tempo, and other features indicated in the score—the author of this manuscript independently performed a detailed auditory evaluation of selected excerpts from publicly available musical literature.

The following works were selected from audio CDs, released by professional record labels, for the auditory scene analysis in the context of the occurrence of perceptual streams:

- Camille Saint-Saëns, *Pour l'indépendance des doigts*, op. 52, no. 2, *Andantino malinconico*, m. 1–14; *Prélude et fugue, F Minor*, op. 52, no. 3, *Fugue, Animato*, m. 1–4; *Prélude et fugue, A Major*, op. 52, no. 5, *Allegro moderato*, m. 1–5, 9–10, performed by: François-René Duchâble, from the record: *Saint-Saëns—6 Etudes Op. 52, Dukas—Sonate, La Plainte Au Loin Du Faune, Prélude Elégiaque*, CDM 7 63159 2, record label: EMI Studio, 1989.
- Franz Liszt, *Réminiscences de Don Juan*, S. 418, Andantino, m. 61–68, performed by: Masaru Okada, from the record: *Cliburn 2001—Semifinal round*, record label: Cliburn, 2001, Recorded live at the Eleventh Van Cliburn International Piano Competition.
- Felix Mendelssohn Bartholdy, *Sechs Lieder ohne Worte*, op. 19, no. 1, *Andante con moto*, m. 3–19, performed by Peter Arne Rohde, from the record: *Mendelssohn—Lieder ohne Worte. Songs without Words*, 1239168, record label: DDD, 1989.

6. Own Research—Sheet Music Analysis, Proving the Occurrence of Multivariantism of Auditory Perceptions

Figure 11 presents an excerpt from Camille Saint-Saëns' piece, *Pour l'indépendance des doigts*, Op. 52, no. 2, *Andantino malinconico*, m. 1–14. Depending on the listener's interpretations related to the incoming sounds, these stimuli can be analyzed using different criteria:

- Variant No. 1: the perceptor hears the melody according to the musical notation as the first perceptual stream (the accented notes, marked as larger in the score), while the second stream is formed by the notes of the harmony, which are also perceived in accordance with the presented score. The third perceptual stream is built from the dotted half notes performed in the instrumentalist's left-hand part in bars 5–8.
- Variant No. 2: the highest notes of the right-hand part form the melody (the first stream), while the second perceptual stream consists of the remaining notes performed in the parts of both hands, forming the harmonic background. The third perceptual stream is formed identically as described in variant No. 1.
- Variant No. 3: "random" sounds chosen by the listener integrate into one perceptual stream, forming the melodic line of the piece. In this context, it is very difficult to refer to specific sounds as the building blocks of the said melody, as it depends on the individual characteristics of the listener as well as the performance method (interpretation) chosen by the instrumentalist. In this case, the shape and direction of the melodic line cannot be described since the frequently repeated sounds that appear in harmonic structures overlap and mask each other (Bremen and Middlebrooks 2013; Huron 2016). Thus, the sound picture is not uniform or clear. The harmonic clusters sounding at the same time cause difficulties in the accurate perception of individual sounds, which might be the reason for the formation of various auditory illusions in the perceptor's mind. The second stream consists of the remaining sounds from the harmonic structures that have not been associated with the melody of the piece, while the third perceptual stream is formed according to the description of variant No. 1.
- Variant No. 4 applies only to the bass line. In this variant, the shape of the melodic line, as well as the harmony, can be a combination of any of the variants discussed above. However, in addition to the previously formed streams, the melody written in the bass part, due to its low register, can be separated by the listener from the other sounds. The formation of this new stream is especially visible in measures 5–8, but it

is also possible for the perceptor to notice this stream earlier and observe its melody until measure 14.

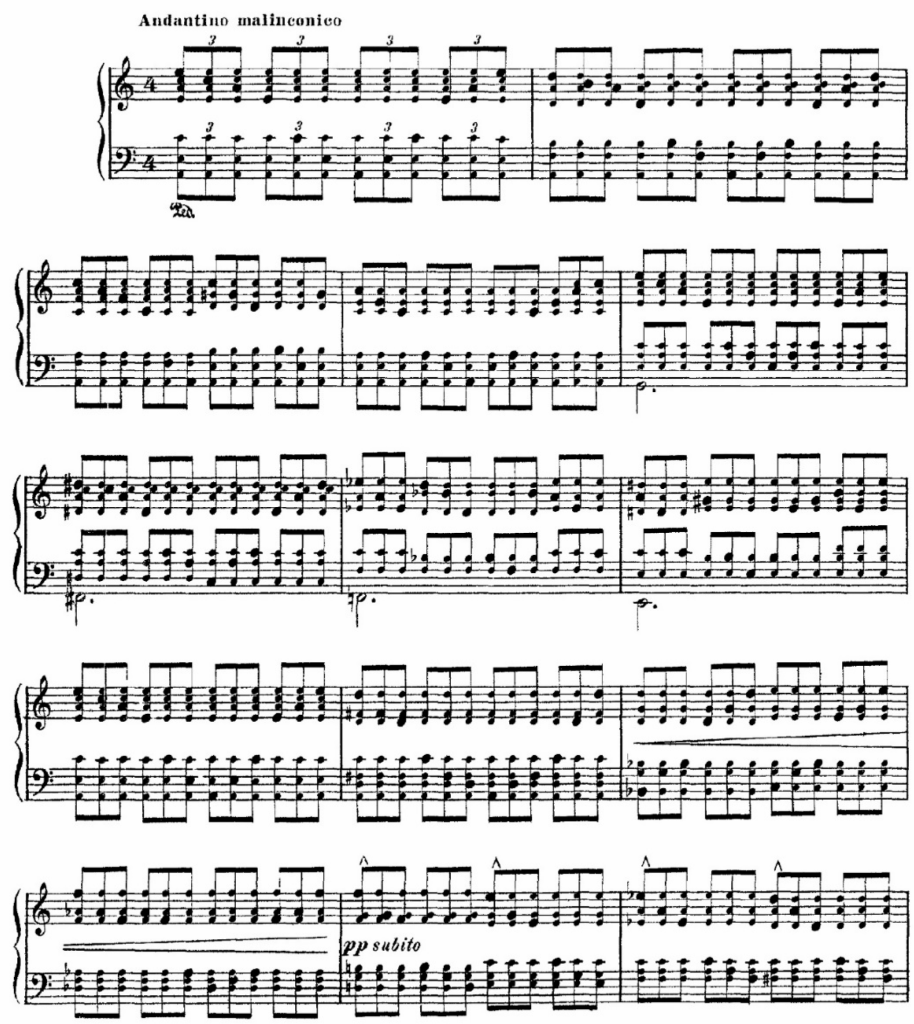

Figure 11. Camille Saint-Saëns, *Pour l'indépendance des doigts*, op. 52, no. 2, *Andantino malinconico*, m. 1–14. Source: https://imslp.org/wiki/Special:IMSLPImageHandler/27025/cy28 (accessed on 23 August 2023).

Figure 12 presents an excerpt from Camille Saint-Saëns's piece *Prélude et fugue, F Minor*, op. 52, no. 3, *Fugue, Animato*, m. 1–4. The sounds in these measures can be perceived in two different ways:

- Variant No. 1: all sounds are integrated into a single perceptual stream.
- Variant No. 2: the two eighth notes written in the bass clef in measures 1 and 2 may not be assigned to the first perceptual stream—the melodic line. This kind of perception occurs when the Ab3, G3, Gb3, and F3 notes are accented more heavily by the performer. The mentioned sounds will then be disconnected from one stream of the melodic line and will be treated as completely "separate", not assigned to any stream. The reason for an interpretation like this is that there are too few sounds (only four) to make the creation of another (second) perceptual stream possible.

Figure 12. Camille Saint-Saëns, *Prélude et fugue, F Minor*, op. 52, no. 3, *Fugue, Animato*, m. 1–4. Source: https://imslp.org/wiki/Special:IMSLPImageHandler/27025/cy28 (accessed on 23 August 2023).

Camille Saint-Saëns's piece, *Prélude et fugue, A Major*, op. 52, no. 5, *Allegro moderato*, m. 1–5 presented in Figure 13 may also be perceived using different criteria:

- Variant No. 1: the sounds written in the right-hand part are integrated into the first perceptual stream, while the sounds from the left-hand part are integrated into the second one.
- Variant No. 2: the highest sounds of the harmonic passages—C#6 for measures 1 to 4 and D6, C#6 and B5 in measure 5 can form the first perceptual stream. The second stream consists of the remaining sounds performed in the part of the right hand, while the third stream is formed from the sounds performed by the instrumentalist with the left hand.
- Variant No. 3 is not an expected variant. However, it can be perceived by a very careful analyst. The dyads played by the instrumentalist in the right-hand part can split so that the listener will perceive the higher sounds as belonging to one stream, while the lower ones will form the second perceptual stream in the listener's mind. The third perceptual stream is formed in the same way as described in variant No. 2.

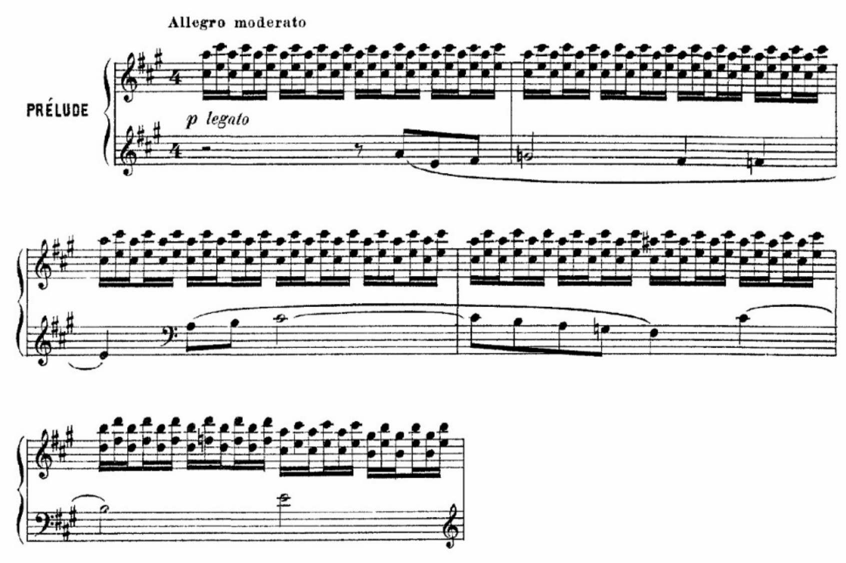

Figure 13. Camille Saint-Saëns, *Prélude et fugue, A Major*, op. 52, no. 5, *Allegro moderato*, m. 1–5. Source: https://imslp.org/wiki/Special:IMSLPImageHandler/27025/cy28 (accessed on 23 August 2023).

Figure 14 shows a further section of *Allegro moderato* (measures 9 and 10), which may be perceived in two different ways:

- Variant No. 1: the sounds performed by the left and right hands are segregated into two perceptual streams. In this case, an inquisitive listener—whether a person with musical training, a performer well-acquainted with the parts they are hearing, or someone meticulously analyzing each sound—will separate the sounds produced by the two hands by observing the differences in pitch between the musical material in the left and right-hand parts.
- Variant No. 2: all the occurring sounds are integrated into a single perceptual stream. Due to the identical rhythm, voice leading, and accents, the incoming sounds merge

into one perceptual stream in the listener's mind. The reason for such an interpretation is the sounds occurring at the same time which can mask each other (Bregman 1990; Huron 2016), making it very difficult for the listener to separate individual tones and associate them with two different streams. Additional features contributing to the integration are the quick succession of sounds (high tempo) and the small subdivision—sixteenth notes.

Figure 14. Camille Saint-Saëns, *Prélude et fugue, A Major*, op. 52, no. 5, *Allegro moderato*, m. 9–10. Source: https://imslp.org/wiki/Special:IMSLPImageHandler/27025/cy28 (accessed on 23 August 2023).

Figure 15 presents an excerpt from Franz Liszt's piece, *Réminiscences de Don Juan*, S. 418, *Andantino*, m. 61–68, which can be interpreted in various ways:

- Variant No. 1: all sounds are integrated into a single perceptual stream, which is the most common observation in this section of the mentioned piece. The factors contributing to the integration are the proximity of the voices on the pitch scale (left- and right-hand parts), as well as the nearly identical rhythm of the voices (except for measures 61, 62, and 68), which results in a synchronized attack and decay of the sounds, causing them to mask each other (Ekström and Borg 2011; McAdams 2024).
- Variant No. 2: very difficult for the perceptor to grasp. However, it may appear in the case of individuals who are able to analyze the incoming sounds very accurately. The sounds perceived by the listener may split into two perceptual streams, especially in bars 64, 66, and 67, when a greater ambitus can be observed between the sounds of the left- and right-hand parts. Increasing the frequency separation between certain parts can be the reason for the segregation of sounds into two perceptual streams (Norman 1966). The segregation phenomenon is most likely to occur for the listeners who are very familiar with the mentioned piece or for the instrumentalists who have created their own interpretation of the piece during the performance.

Figure 15. Franz Liszt, *Réminiscences de Don Juan*, S. 418, *Andantino*, m. 61–68. Source: https://ks15.imslp.org/files/imglnks/usimg/e/e5/IMSLP95879-PMLP08541-Liszt_Klavierwerke_Peters_Sauer_Band_8_14_Don_Juan_Phantasie_scan.pdf (accessed on 6 December 2024).

Figure 16 shows an excerpt from Felix Mendelssohn Bartholdy's piece, *Sechs Lieder ohne Worte*, Op. 19, no. 1, *Andante con moto*, m. 3–19. This section is very perceptually interesting due to the occurring inconsistencies in the reception of the auditory image:

- Variant No. 1: the listener perceives two streams. One of them consists of the highest sounds of the right-hand part, which form the melody, while the second perceptual stream is made of all the other notes that build harmony in the perceptor's mind.
- Variant No. 2: in this case, the listener also perceives only two perceptual streams. The highest sounds from the part of the right hand and the lowest from the part of the left hand are integrated into one perceptual stream, while the other is created on the basis of sixteenth-note passages performed "in the middle"—between the highest and lowest sounds from the parts of both hands.
- Variant No. 3: the material is separated into three perceptual streams. The first stream consists of the highest notes written in the right-hand part (as in variant 1) and makes up the melody in the mind of the listener. The second perceptual stream is formed on the basis of the lowest tones performed with the left hand, which produces the second melodic line. The third and last perceptual stream is formed (as in variant 2) by the sounds that are "in the middle", between the highest and lowest sounds in the parts of both hands—written as sixteenth-note passages.
- Variant No. 4: involves only two perceptual streams, yet its interpretation may be difficult to grasp. A melody written in a treble clef combined with two (measures 3–11) or three (measures 11–14) sixteenth notes form the first perceptual stream. The sounds of the harmony (sixteenth notes) mask the highest notes (the melody) (Huron 2016), so the listener perceives the image of the sounds of the melody rising above the sounds of the harmony but attributes the heard phenomenon to a single stream. The second perceptual stream contains the sounds assigned to the part of the instrumentalist's left hand. This stream is formed by the lowest notes performed in the left-hand part, along with one or two sixteenth notes assigned to this stream based on pitch proximity. Just like in the first stream, the sound of the melody in the left-hand part rises above the sixteenth-note passages and is well heard by the listener.

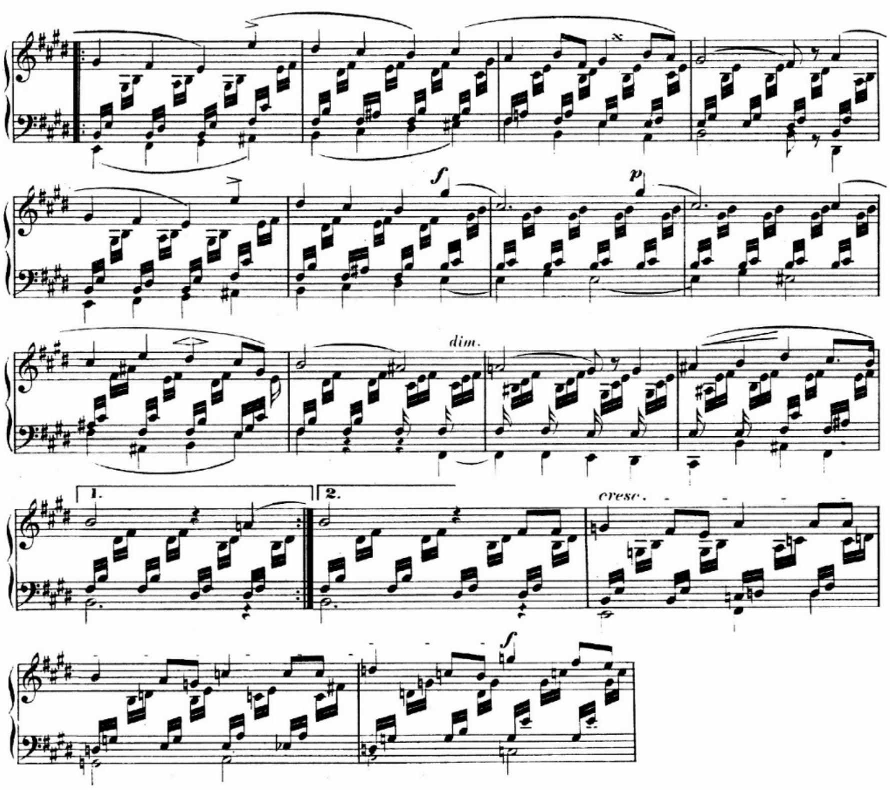

Figure 16. Felix Mendelssohn Bartholdy, *Sechs Lieder ohne Worte*, op. 19, no. 1, *Andante con moto*, m. 3–19. Source: https://imslp.org/wiki/Special:ImagefromIndex/52241/cy28 (accessed on 2 November 2023).

7. Conclusions

It is worth remembering that music is not just sound but also emotions and spiritual experiences conveyed in the form of an interpretation. The coexistence of many sounds that overlap and interfere with one another, contributing to a whole, which makes up the outstanding sound values of a given musical work, etc., causes a variety of perceptual sensations. If sound is considered solely within the frameworks of physics (acoustics) or psychology, it may lack immediate artistic value, as the technical language of these fields often diverges from the expressive language of music (Cf. Rosiński 2020; Adamo 2011). However, interpreting and visualizing acoustic signals are not only crucial for real-world acoustic assessments but also offer valuable insights that enhance our understanding of musical phenomena. By decoding patterns and structures within acoustic data, these scientific methods offer valuable perspectives that can enhance artistic interpretation and foster a deeper connection between scientific and musical domains. Therefore, adopting a multidirectional approach that incorporates both artistic sensibilities and scientific rigor is essential to fully appreciate and understand a given piece. Such an approach allows for a comprehensive grasp of music, revealing layers of meaning that might remain hidden when viewed from a single disciplinary lens.

The highly analytical nature of auditory perception brings to light the phenomenon and the incredible complexity of the hearing process. The auditory image analysis proves how many different perceptual variants can be observed depending on the context. The knowledge of this subject is important for musicians, as it allows them to look at the auditory scene analysis in a completely different way. It turns out that the variety of interpretations of the same sound stimuli provides some previously unrecognized potential of the listeners, which should be thoroughly studied and analyzed. The formation of different mental figures in subjects' minds when stimulating their auditory organ with the same sound wave should not be understood by researchers as an incorrect recognition of the acoustic stimulus but rather as some kind of a unique ability of the listeners, which can reveal itself in certain perceptual situations.

The presented method of auditory analysis proves how much is still to be discovered and understood in this field. Researchers often explore the question of how we perceive various sound structures. Certain transience of observations or thoughts is very characteristic for music, as this issue does not occur to such a wide extent, for example, in fine arts. The impressions that a listener gets from listening to music can be at least partially understood after elementary study of the reception of perceptual streams, which, under various circumstances and depending on different conditions of each listener (focus, order of interpretation of musical piece components at the given moment, etc.) (Bregman 1990) can form various musical structures (Bregman and Ahad 1996).

The occurrence of perceptual streams in musical pieces proves that besides the aspects undertaken by music theory understood in the traditional way, there are some other aspects deciding on the reception of sounds. These possibilities have not yet been fully studied and understood; hence, we do not have sufficient knowledge in this field. Therefore, it is worth exploring it in order to learn about the causes of the phenomena that direct the human brain during perception. It should be noted that the new approach does not depreciate or detract from the previous research achievements in this area but rather complements them. Music theory, like any other field, is constantly developing, so it is worth considering the occurrence of perceptual streams in musical works, as they determine the reception of specific works and thus influence their interpretations.

Interfering with sound material through in-depth auditory scene analysis lets us discover various psychoacoustic phenomena accompanying the listeners, regardless of their chosen act of cognition. The research of this matter may head toward a significant widening of the cognition boundaries and help to deepen the knowledge about human auditory skills and abilities, revealing themselves during the reception of musical pieces. The described perceptual phenomena, when introduced to the field of music theory, allow us to explore musical works from a new, previously undiscovered research perspective.

8. Author's Commentary

Psychoacoustics, the science of human sound perception, is an interdisciplinary field encompassing a variety of research approaches and applications. Its scope and uses can be categorized into five main perspectives, which, although unified by a common focus on sound perception, vary significantly in their research subjects, methodologies, and scientific goals.

- Musical Art: Music Perception and Analysis of Musical Reception (the focus of this text)

The first perspective in psychoacoustics pertains to the study of music perception, analyzing how individuals perceive and interpret music. In this context, psychoacoustic research focuses on the structural and perceptual analysis of musical phrases, chords, harmonies, and rhythms, exploring the auditory and emotional responses these elements evoke in listeners. Methodologically, this resembles musical analysis, wherein a theorist examines the components of a piece as they would study a musical score, forming conclusions and insights. Musical psychoacoustics frequently draws on music theory and cognitive science to understand how listeners discern specific patterns and structures, as well as how these perceptual structures affect the reception of music. The findings from such research can lead to a deeper understanding of aesthetic elements and the influence of cultural and emotional context on music perception.

- Psychology: Studies of Sound Perception on Large Research Samples

The second perspective in psychoacoustics draws on psychology to study sound perception across large, representative groups, producing statistically significant results. This research typically involves simple tones—sound signals with regular acoustic properties that are easy to measure and control in laboratory settings. The use of simple tones avoids the influence of complex musical structures or the varying timbers of different instruments, allowing researchers to focus on universal principles of sound perception, such as auditory masking, thresholds of audibility, and spatial localization. While these studies provide fundamental knowledge about perceptual mechanisms, they do not delve into the complexities of musical analysis.

- Physics: Applications of Sound in Environmental and Spatial Acoustics Research

The third perspective applies psychoacoustics to environmental research, particularly in the context of environmental and spatial acoustics. Sound is used to analyze the physical properties of environments, such as the depth and structure of seabeds (e.g., through sonar technology), as well as to study sound propagation in air or various materials. In noise studies, psychoacoustics assesses the impact of noise on surroundings, playing a crucial role in designing public spaces and protecting the environment from noise pollution. Additionally, this perspective applies to architectural acoustics, where it helps optimize the acoustics of concert halls, recording studios, and other spaces with specific acoustic requirements. This approach focuses primarily on practical and technical applications, aiming to ensure appropriate acoustic conditions and minimize the harmful effects of noise.

- Sound Engineering: Studies of Electroacoustic Devices and Their Impact on Hearing

In sound engineering, psychoacoustics examines the properties of electroacoustic devices, such as speakers, microphones, and amplifiers, and their effects on human sound perception. These studies often use artificially generated sounds, allowing precise control over acoustic parameters like frequency, amplitude, and phase. Psychoacoustics helps optimize audio technologies to maximize sound quality and minimize potentially harmful effects on hearing, such as distortion or excessive sound pressure levels. Applications in this domain include the development of modern audio systems, such as surround-sound technologies and automotive acoustic designs, as well as the creation of headphones with active noise cancellation.

- Medicine: Studies of Neural Processing and Hearing Disorders

The final psychoacoustic perspective focuses on medicine, particularly in audiology and neurology, where research investigates the mechanisms of sound processing in the nervous system and associated disorders. Medical psychoacoustics examines how sound is processed by various components of the auditory system and identifies factors that may disrupt this process, such as injury, aging, or genetic conditions. This perspective also includes the study of speech disorders, hearing impairments, and other conditions related to sound perception, which can lead to communication difficulties and a reduced quality of life. The goal of this research is to develop therapeutic and diagnostic methods for the early detection and treatment of hearing problems, as well as to support patient rehabilitation, for instance, through hearing aids or cochlear implants.

- Summary: Music as a Distinct Domain in Psychoacoustics

Although music is one of the subjects of psychoacoustic research, it represents a distinct field that differs from the other perspectives discussed above. The study of music perception involves unique qualities of complex sounds and the analysis of aesthetic and emotional aspects, distinguishing it from research in psychology, physics, engineering, or medicine. As an art form, music encompasses a richness of tonal and harmonic structures that deeply influence emotional and intellectual reception, presenting music psychoacoustics with challenges unlike those faced by researchers analyzing simple tones, noise, or acoustic devices. This text is primarily directed at musical artists, who, although regularly exposed to psychoacoustic phenomena, are often unaware of the subtle perceptual mechanisms shaping their music perception and interpretation. Understanding how complex psychoacoustic phenomena influence emotions, aesthetic impressions, and the interpretation of musical structures can enhance artistic sensitivity, enabling a more intentional and expressive use of sound.

Funding: This research was funded by the Faculty of Art at the University of Warmia and Mazury in Olsztyn. The APC was funded by the Faculty Research Fund of the Faculty of Art allocated for the year 2024.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The author declares no conflict of interest.

References

Adamo, Giorgio. 2011. Single Tone Colour and the ‘Sound’ of Music Items. Some Remarks on the Qualitative Aspects in Music Perception. In *Klangfarbe. Vergleichend-Systematische und Musikhistorische Perspektiven*. Edited by Alois Schmidhofer and Stefan Jena. Frankfurt am Main: Peter Lang, pp. 143–58.

Barrett, Karen, Richard Ashley, Diana Strait, Erika Skoe, Charles Limb, and Nina Kraus. 2021. Multi-Voiced Music Bypasses Attentional Limitations in the Brain. *Frontiers in Neuroscience* 15: 1–7. [CrossRef] [PubMed]

Bigand, Emmanuel, Stephen McAdams, and S. Forêt. 2000. Divided Attention in Music. *International Journal of Psychology* 35: 270–8. [CrossRef]

Botte, Marie-Claire, Carolyn Drake, Renaud Brochard, and Stephen McAdams. 1997. Perceptual Attenuation of Nonfocused Auditory Streams. *Perception & Psychophysics* 5: 419–20. [CrossRef]

Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: The MIT Press.

Bregman, Albert S., and Pierre Ahad. 1996. Audio Demonstrations of Auditory Scene Analysis. *Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound.* Available online: https://themusiclab.github.io/bregman-archive/downloadsdl.htm (accessed on 15 August 2024).

Bremen, Peter, and John C. Middlebrooks. 2013. Weighting of Spatial and Spectro-Temporal Cues for Auditory Scene Analysis by Human Listeners. *PLoS ONE* 8: e59815. [CrossRef]

Brewer, William F., and Lester Loschky. 2005. Top-Down and Bottom-Up Influences on Observation: Evidence from Cognitive Psychology and the History of Science. In *Cognitive Penetrability of Perception: Attention, Action, Strategies, and Bottom-Up Constraints*. Edited by Athanassios Raftopoulos. New York: Nova Science, pp. 31–47.

Carey, Daniel, Stuart Rosen, Saloni Krishnan, Marcus T. Pearce, Alex Shepherd, Jennifer Aydelott, and Frederic Dick. 2015. Generality and Specificity in the Effects of Musical Expertise on Perception and Cognition. *Cognition* 137: 81–105. [CrossRef]

De Poli, Giovanni. 2004. Models for Musical Analysis. In *Music Information Processing*. Edited by Giovanni De Poli. Padova: Università degli Studi di Padova.

Devinsky, Orrin, and Mark D'esposito. 2004. *Neurology of Cognitive and Behavioral Disorders*. New York: Oxford University Press.

Disbergen, Niels R., Giancarlo Valente, Elia Formisano, and Robert J. Zatorre. 2018. Assessing Top-Down And Bottom-Up Contributions to Auditory Stream Segregation and Integration with Polyphonic Music. *Frontiers in Neuroscience* 12: 121. [CrossRef]

Dykes, Brent. 2020. *Effective Data Storytelling. How to Drive Change with Data, Narrative and Visuals*. Hoboken: Wiley.

Ekström, Seth-Reino, and Erik Borg. 2011. Hearing Speech in Music. *Noise & Health* 13: 277–85. [CrossRef]

Hove, Michael, Celine Marie, Ian C. Bruce, and Laurel J. Trainor. 2014. Superior Time Perception for Lower Musical Pitch Explains why Bass-Ranged Instruments Lay Down Musical Rhythms. *Proceedings of the National Academy of Sciences* 111: 10383–8. [CrossRef]

Hubbart, Timothy L. 2010. Auditory Imagery: Empirical Findings. *Psychological Bulletin* 136: 302–29. [CrossRef]

Huron, David. 1991. Tonal Consonance Versus Tonal Fusion in Polyphonic Sonorities. *Music Perception: An Interdisciplinary Journal* 9: 135–54. [CrossRef]

Huron, David. 2016. *Voice Leading. The Science Behind a Musical Art*. Cambridge, MA: The MIT Press. [CrossRef]

Jain, Saransh, and Nuggehalli Puttareviyah Nataraja. 2019. The Effect of Fatigue on Working Memory and Auditory Perceptual Abilities in Trained Musicians. *American Journal of Audiology* 28: 483–94. [CrossRef] [PubMed]

Jain, Saransh, Nuggehalli Puttareviyah Nataraja, and Vijaya Kumar Narne. 2022. The Effect of Subjective Fatigue on Auditory Processing in Musicians and Nonmusicians. *Music Perception* 39: 309–19. [CrossRef]

Janata, Petr, Barbara Tillmann, and Jamshed J. Bharucha. 2002. Listening to Polyphonic Music Recruits Domain-General Attention and Working Memory Circuits. *Cognitive, Affective, & Behavioral Neuroscience* 2: 121–40. [CrossRef]

Kashino, Kunio. 2006. Auditory Scene Analysis in Music Signals. In *Signal Processing Methods for Music Transcription*. Edited by Anssi Klapuri and Manuel Davy. Boston: Springer, pp. 299–325. [CrossRef]

Keller, Andreas. 2016. *Philosophy of Olfactory Perception*. Cham: Palgrave Macmillan. [CrossRef]

Kin, Maurycy J., and Andrzej Dobrucki. 2016. Perception of Changes in Spectrum and Envelope of Musical Signals vs. Auditory Fatigue. *Archives of Acoustics* 41: 323–30. [CrossRef]

Kirwan, Brock. 2016. Cognitive Neuroscience. In *The SAGE Encyclopedia of Theory in Psychology*. Edited by Harold L. Miller, Jr. Los Angeles: Sage, vol. 1. [CrossRef]

Koelsch, Stefan. 2011. Toward a Neural Basis of Music Perception—A Review and Updated Model. *Frontiers in Psychology* 2: 110. [CrossRef]

Kondo, Hirohito, and Makio Kashino. 2009. Involvement of the Thalamocortical Loop in the Spontaneous Switching of Percepts in Auditory Streaming. *Journal of Neuroscience* 29: 12695–701. [CrossRef]

Kornmeier, Jürgen, and Michael Bach. 2012. Ambiguous figures—What happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience* 6: 51. [CrossRef]

Macpherson, Fiona. 2006. Ambiguous Figures and the Content of Experience. *NOÛS* 40: 82–117. [CrossRef]

McAdams, Stephen. 2024. *Perception and Cognition of Music. The Sorbonne Lectures*. New York: Oxford University Press.

McAdams, Stephen, and Albert Bregman. 1979. Hearing Musical Streams. *Computer Music Journal* 3: 26–43+60.

McAdams, Stephen, Philippe Depalle, and Eric Clarke. 2004. Analyzing Musical Sound. In *Empirical Musicology: Aims, Methods, Prospects*. Edited by Eric Clarke and Nicholas Cook. New York: Oxford Academic Press, pp. 157–96. [CrossRef]

McKenna, Eugene. 2020. *Business Psychology and Organizational Behaviour*, 6th ed. New York: Routledge. [CrossRef]

Moore, Travis M., Alexandra P. Key, Antonia Thelen, and Benjamin W.Y. Hornsby. 2017. Neural mechanisms of mental fatigue elicited by sustained auditory processing. *Neuropsychologia* 106: 371–82. [CrossRef]

Norman, Donald A. 1966. *Rhythmic Fission: Observations on Attention, Temporal Judgments and Critical Band*. Cambridge: Harvard University, Unpublished manuscript.

Palmer, Caroline. 2015. Listening, Imagining, Performing: Melody as a Life Cycle of Musical Thought. *Music Perception* 33: 3–11. [CrossRef]

Pearson, Joel. 2019. The Human Imagination: The Cognitive Neuroscience of Visual Mental Imagery. *Nature Reviews Neuroscience* 20: 624–34. [CrossRef] [PubMed]

Pecenka, Nadine, and Peter E. Keller. 2009. Auditory Pitch Imagery and Its Relationship to Musical Synchronization. *Annals of the New York Academy of Sciences* 1169: 282–6. [CrossRef]

Poeppel, David, and Tobias Overath. 2012. Introduction: Why Human Auditory Cortex. In *The Human Auditory Cortex*. Edited by David Poeppel, Tobias Overath, Arthur N. Popper and Richard R. Fay. Springer Handbook of Auditory Research. Editors of series Richard R. Fay and Arthur N. Popper. New York, Dordrecht, Heidelberg and London: Springer. [CrossRef]

Pressnitzer, Daniel, Clara Suied, and Shihab A. Shamma. 2011. Auditory Scene Analysis: The Sweet Music of Ambiguity. *Frontiers in Human Neuroscience* 5: 158. [CrossRef] [PubMed]

Rajendran, Vani G., Sundeep Teki, and Jan W.H. Schnupp. 2018. Temporal Processing in Audition: Insights From Music. *Neuroscience* 318: 4–18. [CrossRef] [PubMed]

Repp, Bruno H. 2009. Segregated in Perception, Integrated for Action: Immunity of Rhythmic Sensorimotor Coordination to Auditory Stream Segregation. *The Quarterly Journal of Experimental Psychology* 62: 426–7. [CrossRef] [PubMed]

Rosiński, Adam. 2020. Perception of Sound via Auditory Image Analysis, Made in the Conceptual Context of the Philosophy of Music. In *Humanitarian Corpus*. Kyiv: M.P. Drahomanov National University of Pedagogy Academic Press, Vinnytsia: TVORY, vol. 2, issue 35.

Rosiński, Adam. 2021. Wpływ wykształcenia muzycznego na grupowanie dźwięków sekwencji ABA-ABA w rytm galopujący. In *Przestrzenie Akustyki. Professional Acoustics*. Edited by Adam Rosiński. Olsztyn: Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, pp. 53–70.

Rosiński, Adam. 2023. Influence of Music Education and Pitch Scales on the Grouping of the AB-AB Sequence Sounds. In *Przestrzenie akustyki. Professional Acoustics*. Edited by Adam Rosiński. Olsztyn: Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, vol. 2, pp. 53–74.

Rosiński, Adam. 2024. *Wpływ melodyki i rytmiki na tworzenie wielowariantowości odbioru wrażeń słuchowych. Przetwarzanie dźwięków w strumienie percepcyjne na przykładzie wybranych dzieł muzycznych*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego. [CrossRef]

Russ, John C., and F. Brent Neal. 2016. *The Image Processing Handbook*, 7th ed. Boca Raton: CRC Press. [CrossRef]

Salamon, Justin, Emilia Gómez, and Joan Serrà. 2012. Melody, Bass Line, and Harmony Representations for Music Version Identification. In *Proceedings of the 21st International World Wide Web Conference (WWW 2012): 4th International Workshop on Advances in Music Information Research (AdMIRe 2012)*. New York: Association for Computing Machinery, pp. 887–94. [CrossRef]

Scheffer, David, and Björn Manke. 2018. The Significance of Implicit Personality Systems and Implicit Testing: Perspectives From PSI Theory. In *Why People Do the Things They Do. Building on Julius Kuhl's Contributions to the Psychology of Motivation and Volition*. Edited by Nicola Baumann, Miguel Kazén, Markus Quirin and Sander L. Koole. Göttingen: Hogrefe, pp. 281–300. [CrossRef]

Siedenburg, Kai, Kirsten Goldmann, and Steven Par. 2021. Tracking Musical Voices in Bach's the Art of the Fugue: Timbral Heterogeneity Differentially Affects Younger Normal-Hearing Listeners and Older Hearing-Aid Users. *Frontiers in Psychology* 12: 608684. [CrossRef] [PubMed]

Simon, Júlia, Endre Takács, Gábor Orosz, Borbála Berki, and István Winkler. 2020. Short-term cognitive fatigue effect on auditory temporal order judgments. *Experimental Brain Research* 238: 305–19. [CrossRef]

Spada, Danilo, Laura Verga, Antonella Iadanza, Marco Tettamanti, and Daniela Perani. 2014. The Auditory Scene: An fMRI Study on Melody and Accompaniment in Professional Pianists. *NeuroImage* 102: 764–75. [CrossRef]

Strumiłło, Paweł. 2012. *Elektroniczne Systemy Nawigacji Osobistej dla Niewidomych i Słabowidzących*. Łódź: Politechnika Łódzka, Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki.

Thompson, Hilary W. 2011. The Biological and Computational Bases of Vision. In *Computational Analysis of the Human Eye with Applications*. Edited by Sumeet Dua, Rajendra U. Acharya and E. Y. K. Ng. Singapore: World Scientific, pp. 1–38. [CrossRef]

Trainor, Laurel J. 2015. The Origins of Music in Auditory Scene Analysis and the Roles of Evolution and Culture in Musical Creation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370: 20140089. [CrossRef]

Trainor, Laurel J. 2021. Understanding the Origins of Musicality Requires Reconstructing the Interactive Dance Between Music-Specific Adaptations, Exaptations, and Cultural Creations. *Behavioral and Brain Sciences* 44: e116. [CrossRef] [PubMed]

van Noorden, Leon P. A. S. 1975. Temporal Coherence in the Perception of Tones Sequences. Ph.D. thesis, Technical University Eindhoven, Eindhoven, The Netherlands.

Wang, DeLiang, and Peter Chang. 2008. An Oscillatory Correlation Model of Auditory Streaming. *Cognitive Neurodynamics* 2: 7–19. [CrossRef] [PubMed]

Wright, James K., and Albert S. Bregman. 1987. Auditory Stream Segregation and the Control of Dissonance in Polyphonic Music. *Contemporary Music Review* 2: 63–92. [CrossRef]

Załazińska, Aneta. 2016. *Obraz, Słowo, Gest*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Zawadzka-Gołosz, Anna. 2018. Gra kompozytora z percepcją słuchacza na przykładzie twórczości Witolda Lutosławskiego. *Res Facta Nova. Teksty o muzyce współczesnej* 19: 61–71. [CrossRef]