sciendo

GESTALT THEORY, DOI 110.2478/gth-2023-0020
© 2023 (ISSN 2519-5808); Vol. 45, No. 3, 183–197

Original Contribution - Originalbeiträge

Nicola Di Stefano

Temporal and spatial accounts of sound perception. An overview of the main historical sources and theoretical problems

1. The temporal nature of sound perception

When we listen to music, we do not listen to music *only*. While my hi-fi stereo system is playing Beethoven's *Leonore* overture, my children are playing "knock down ginger" with neighbours, and a wide variety of other noises – street traffic, police siren, air conditioning fan, rain – simultaneously reach my ears. However, if someone asks me what I am doing, presumably I will answer that I am listening to my favourite Beethoven's overture.

The above example represents a case of what Cherry (1953) defined "cocktail party" effect, referring with these terms to the typical situation we experience when at a party we are able to identify a distinct stream of auditory information, for instance, a person talking to us, from other voices and background noises. More recently, Bregman (1990) introduced the psychological concept of 'auditory scene analysis' to explain the ability to isolate one particular group of sounds from the complexity of the listening scenarios we are exposed to daily.

In the auditory stream, perceptual units emerge from the parsing of the continuous temporal auditory flow into fragments thanks to some kinds of *Gestalt* principles, that allows for distinguishing auditory objects from one another and from the background (Kubovy & Van Valkenburg, 2001). Temporal proximity and similarity have been designated as the two primary factors of the cohesion and segregation of auditory units (see, e.g., Tenney & Polansky, 1980). Proximity means, for example, that different scratching noises heard close together in time will likely be grouped together, while similarity means that a siren sounding every five seconds will be unified as a perceptual unit (see also Griffiths & Warren, 2004).

In this perspective, the auditory object emerges as "the computational result of the auditory system's ability to detect, extract, segregate and group the spectrotemporal regularities in the acoustic environment into stable perceptual units" (Bizley & Cohen, 2013, p. 693). At the same time, however, studies have demonstrated that the concept of regularity is not to be rigidly conceived.

For example, listeners typically ignore small variations in the temporal organization to save patterns of regularities (e.g., von Ehrenfels, 1988). When a sequence of metronome clicks is slightly irregular though still remaining within the tolerance window, then our perception 'adjusts' the irregularities, and we perceive the sequence as the succession of the 'same' temporal intervals (Drake & Bertrand, 2003).

The temporal nature of sounds perception has inspired several psycho-musicological models in which sounds are perceived according to their ability to fulfil expectations within the perceptual flow (e.g., Huron, 2006; Narmour, 1990; Margulis, 2005). Time played a key role also in the conceptualization of sounds proposed by composers, with Stravinsky observing that: «Music is a *chronologic* art, as painting is a *spatial* art. Music presupposes before all else a certain organization in time, a chrononomy» (Stravinsky, 1947, p. 28; see also Schoenberg, 1950). Schaeffer conceived the sonorous object phenomenologically as an *intentional* unit, constituted in our consciousness by our own mental activity (Schaeffer, 1966, p. 263). In this account, as evidenced by Godøy (2021), temporal duration is probably the most relevant criterion for sound object formation, with acoustic features of sound objects ranging from the slow timescale of the shape of the entire sound object.

Finally, the temporal nature of auditory objects has also oriented the philosophical debate, leading researchers to emphasize the opposition between vision and hearing – spatial the former and temporal the latter – and ultimately to claim the alleged primacy of vision in shaping our conceptualization of perceptual objects (Kubovy, 1988; Kubovy & Van Valkenburg, 2001; see also Casati & Dokic, 2004; Casati, Dokic, Di Bona, 2020).

That being said, a caveat should be made regarding the temporal nature of sound perception. As Husserl noted, the temporal nature of the auditory flow is such that it is characterized by retention and protention, i.e., the abilities that extend actual perception back to the past and toward the future, respectively (Husserl, 1991). In this perspective, a pivotal role in determining what is past and future is played by the 'present now'; however, such concept seems blurred, given that the present moment of music has always been already *experienced*, and thus it is part of the past.

Such a paradoxical nature of sound perception was grasped by Hegel, who characterized sound itself as "an externality which in its coming-to-be is annihilated again by its very existence, and it vanishes of itself" (Hegel, 1975, p. 890). Thus, it seems that what is referred to as 'now', rather than being a *datum*, results from a dynamic process which necessarily assumes it as being part of the past. Clarke (2011) suggested that such a pitfall might be avoided by expanding the temporal duration of the present in such a way that *heard* elements of a melody are retained in and form the present. This means that the temporality of a melody can be apprehended only as a simultaneity, an observation that would probably align with Husserl's concept of melody as a whole (Husserl, 1991). However, if this is the case, the temporal dimension of a melody would imply simultaneity, and simultaneity, as Clarke (2011, p. 12) himself observed, is the *spatialization* and therefore the *detemporalization* of time (see also Xenakis, 1988). These remarks suggest that the temporal nature of sounds, when applied to the perception of melody, inevitably leads to admitting space as an alternative conceptualization. Thus, in the following section, I will move on to examining how space has entered the conceptualization of music, from philosophy to musicology.

2. Spatial concepts of sound

The concept of space has been applied to sound perception in at least two different ways. First, by referring to the localization of sound sources and the spatial region occupied by the sounding object (see Casati & Dokic, 1994). For example, Nudds (2009) started his investigation of sound with the question: «Where are sounds and where do we experience them to be?» (Nudds, 2009, p. 69). Similar investigations led philosophers of perception to highlight that sound perception informs us both about the position of the sound source and the quality of the object that might have generated the sound (O'Callaghan, 2009). For example, the sound of an object that falls to the ground and breaks will inform us not only about the localization of the event but also about the dimension of the object. For this reason, Matthen (2010) noted that sounds are not merely located but are *object*-located events.

However, an original notion of musical space has been also developed in the literature across psychology, musicology, and philosophy (Di Stefano, 2022, for a review). In this context, space is essentially conceptualized as an autonomous or independent perceptual environment that is intrinsic to musical sounds and in which sounds can be placed and move about (e.g., McDermott, 1972). Hence, the term 'space' loses some of the essential features that characterize it in the visual domain, such as being physical and three-dimensional. As Zuckerkandl put it: "The space we hear is a space without places" (Zuckerkandl, 1956, p. 276). However, as Lippman noted, sound perception can be still conceived of as a different kind of spatial experience «whose nature is open to question» (Lippman, 1952, p. 112). In what follows, I provide a brief historical overview of some prominent spatial conceptualization of sounds, starting from the pioneering contribution of Hermann von Helmholtz.

2.1 Helmholtz

Helmholtz provided the first thorough investigation of the relationship between space and musical tones. In his treatise *On the sensations of tones*, he noted the perceived analogy between musical scale and space: «It is an essential character of space that at every position within it like bodies can be placed, and like motions can occur. Everything that is possible to happen in one part of space is equally possible in every other part of space and is perceived by us in precisely the same way. This is the case also with the musical scale. Every melodic phrase, every chord, which can be executed at any pitch, can be also executed at any other pitch in such a way that we immediately perceive the characteristic marks of their similarity» (Helmholtz, 1954, p. 370). Such perceptual similarity is seemingly based on the properties that music and space phenomenologically exhibit. As we perceive the motion of physical objects in the empirical space, we (can) perceive sounds changing in pitch as if they were moving from one place to another: the dynamic change of pitch over time «has a readily recognised and unmistakable resemblance to motion in space, and is often metaphorically termed the ascending or descending motion» (Helmholtz, 1954, p. 370).

For Helmholtz, the spatialization of sound is assumed to occur at the very preliminary level of sound processing. According to the tonotopic organization of the cochlea, frequency information is transduced into spatial information, based on the resonant frequency specific to each site along the basilar membrane and on the morphological and mechanical characteristics of the organic tissue, such as thickness, elasticity, and stiffness (see Helmholtz, 1954; Warren, 2013). In this view, temporal simultaneity is discretized over different points in space, with the harmonics of complex sounds being spatialized at different locations along the basilar membrane.

The tonotopic theory allows to explain the phenomenon of beating by observing the fact that portions of the membrane set into vibration by sounds close in frequency overlap each other, thus giving rise to the phenomenon of beats. Beatings thus arise when two sounds hit the membrane within the 'critical band', which can be accordingly conceived in terms of the ear's minimum threshold for successfully analyzing the auditory frequencies, below which the ear is unable to separate two simultaneous frequencies.

Shifting from a physiological to a perceptual level, Helmholtz views geometrical properties of the physical space, such as invariance and resemblance, as capable of establishing the parallel to musical space. Helmholtz points out that listeners immediately perceive the identity of intervals transposed along the musical scale. In contrast, we would not be able to perceive such identity in the domain of colour. As Pesic (2013) puts it: "We can transpose a Beethoven sonata up a half step and still recognize the work as in some sense still the same, yet we cannot likewise "transpose" all the colors of a Rembrandt (say by shifting all reds to orange, orange to yellow, and so forth): Though the basic line and surface contours of the painting remain unchanged, its color harmony cannot be "transposed" and remain recognizably identical. Helmholtz extends the special quality of spatial resemblance that can be seen in related shapes (such as the similar profiles of father and daughter) to the characteristic melodic contour of a certain piece of music, but not to colors" (Pesic, 2013, p. 279; see Spence & Di Stefano, 2022, 2023 on crossmodal matchings and sensory translation between colours and sounds).

The parallel between musical sounds and physical space is based on the analogy, or isomorphism, mediated by motion (Helmholtz, 1954; see also Di Stefano, 2022). Though not itself spatial or extended, music can move in analogy to spatial motion, from which Helmholtz boldly identifies the emotional appeal of music: its virtual motion is felt as emotion precisely because of the deep isomorphism between musical and physical space (although the mediation of movement could hardly be evoked to explain the emotional expressivity of timbre, see Di Stefano, 2023).

However, the analogy between musical space and physical space can be also mediated, for Helmholtz, by a more abstract concept, namely the notion of dimensional space, conceived of as a formal entity mainly characterized by being an n-dimensional manifold. "In the same way we may consider the system of simple tones as an aggregate [a manifold] of two dimensions, if we distinguish only pitch and intensity and leave out of account differences of timbre. This generalization of the idea is well-suited to bring out the distinction between space of three dimensions and other aggregates [manifolds]. We can, as we know from daily experience, compare the vertical distance of two points with the horizontal distance of two others, because we can apply a measure first to the one pair and then to the other." (Helmholtz in Pesic, 2007, p. 53).

After Helmholtz, several philosophers and musicologists have elaborated on spatial accounts of musical sounds. Among the issues discussed in the early literature, are the comparison of musical space to the three-dimensional geometrical space, the spatial character and representation of pitch and pitch relationships, and the spatial features of musical time (see Harley, 1994, and Di Stefano, 2022, for reviews). In what follows, we briefly consider the German tradition and the phenomenological approaches.

2.2 The German tradition

The philosopher and musicologist Ernst Kurth conceived of musical space as essentially manifested in and through movement. He noticed in particular the spatial qualities of pitch, melodic line, contrary and oblique motion, as well as the distance of notes in an interval (*Zwischenraum*; Kurth, 1931). Few decades later, the distinction between various types of space in music was refined in the works of Albert Wellek, who identified three different meanings of music space, namely hearing space (*Gehörraum*), tonal space (*Tonraum*), and the proper musical space (*Musikraum*) (Wellek, 1963). Hearing space is the aurally mediated spatial orientation, an incomplete "image" of the objective and physical space. Tonal space is an unsteady, indistinct structure or ordering schema in three dimensions, two of these corresponding to pitch and temporal order, respectively (see Riedel, 2019). Musical space is a "pure" feeling space idiosyncratic of music, feeding itself from the previous (the tonal space, but also the hearing space), substantially though based on expression of feelings in (absolute) music as such.

Despite the theoretical effort made by both Kurth and Wellek, similar descriptions of musical space remain somewhat obscure. Kurth himself was seemingly aware of such risk, when he states that musical space «is not visible, not touchable, and really hardly conceivable» (Kurth, 1931, p. 119). Thus, he used to refer to the idea of space in music in terms of *space-feeling* rather than space-conception. However, conceptual obscurity is no argument against the existence of musical space, but merely against its identity with external space, which rests on the clearest perceptual and intellectual understanding (see Kurth, 1931, p. 127).

A three-dimensional model of musical space (with the axes of pitch, time, dynamics) has been also proposed by the German musicologist and composer Hans-Joachim Moser (1953). This model includes a temporal dimension on one axis and standard dynamic levels (*p*, *mf*, *f*) as units on the other axis.

2.3 Phenomenological approaches: Lippman, McDermott and Clifton

In his doctoral dissertation on music and space at Columbia University, the pioneering American philosopher Edward Lippman explicitly moves from the apparent contradictory nature of the conceptualization of spatiality in music: «The explanation of the spatial aspects of musical experience is obviously not to be found directly in perceptual or empirical space, but there is apparently no meaningful concept of space other than this» (Lippman, 1952, p. 135). Lippman evidenced that the intrinsic spatiality of music is phenomenologically evident, and grounded on structural similarity, i.e., the many identical formal elements that characterize the experiences of music and of empirical space. For Lippman, however, when it comes to analytically describe musical space, the use of concepts elaborated for describing empirical space unavoidably generates the conceptual complexity (if not vagueness): «The tendency of musical space to seek embodiment in conceptions of empirical space is so readily understandable as not to require explanation; an identity of structure and of sensational character, and the dominance of utility in establishing meanings are the chief underlying reasons»

(Lippman, 1952, p. 235). Reflecting on the non-literal use of the term 'space' in similar contexts, thus, Lippman (1952) observed that such musical space is rather a *pseudo-space*.

Key reflections on the notion of musical space were provided by Vincent McDermott and Thomas Clifton. McDermott introduced the notion of musical space as a conceptual structure that allows us to individualize objects and distinguish them from one another. According to him, pitches sounding simultaneously that do not blend together indissolubly, maintain separate *positions*, i.e., they occupy different places, in nothing other than a musical space (McDermott, 1972, p. 490). McDermott grounded his notion of musical space on pitch and depth or masking. The latter dimension is introduced to account for the fact that pitches are perceived as moving not only according to vertical direction, i.e., high-low scale, but also in depth: «Pitches often do not appear on a single plane in our mind's image of the piece, as it grows and as we comprehend it. On the contrary, one sound or group of sounds tends to stand out, to demand more of our attention. Other sounds recede, become obscured by still other sounds» (McDermott, 1972, p. 492). McDermott referred to this perceptual phenomenon with the psychoacoustical concept of masking. In his view, space is a conceptual tool that we implicitly use when listening to music we are aware of relations of height, interval, depth, or counterpoint. Such relations are, according to McDermott, all conceived of as spatial. However, when it comes to the articulation of such musical space and the positioning of sounds, McDermott explanations become much elusive (as highlighted above for Kurth and Wellek): «Every pitch, timbre, dynamic, every group of tones, every formal intricacy, every durational emphasis, even every rest - in sum, everything about a piece of music - contribute in some manner, substantially or only slightly, to the spatial organization of the work» (McDermott, 1972, p. 491).

Delving into the phenomenology of sounds, Cliftons' study of «music as heard» touches upon various aspects of the perceptual experience including musical space. For Clifton, musical space is one of the four essential features of musical experience, together with time, play and feeling and understanding. Such notion of musical space is rooted in the peculiar experience of hearing tones as «occupying certain positions of a purely phenomenal, nonphysical nature» (Clifton, 1983, p. 143). Therefore, the musical space has nothing to share with listener's or sound's localization: «To be in musical space means more than mere existence at a particular place, and therefore has nothing to do with one's physical location» (Clifton, 1983, p. 141). The notion of space embraces diverse aspects of the perceptual experience of music, for example, line and surface among others. Line is defined as «the narrowest, if not the simplest form of musical space is the single line» (Clifton, 1983, p. 143). The musical line exhibits some of the properties of geometrical lines, for instance, thickness. Whilst the association of melody with the concept of line seems to be intuitive (e.g., see the common description of melodies as melodic lines), Clifton's characterization of surfaces as elements that vary according to several musical parameters, such as dynamics, intensity, timbral complexity (Clifton, 1983, p. 155) seems vague. In this respect, I must agree with Macedo (2015), who observed that Clifton's notions of line and surface seem to be unclear and unclearly differentiated. In fact, although they seem to be modeled after empirical space, their distinction does not seem to correspond to a similar distinction in empirical space. Therefore, the impression is that the use of geometrical/visual terms in musical context is exclusively based on metaphorical mapping, that nevertheless remains unclear, and thus fails in clarifying how space can be conceived of as an organization principle for sounds perception.

3. Interim summary

The various spatial accounts of sounds summarized here (see, also, Morgan, 1980, and Zbikowski, 2002) have a certain number of common features. First, they assume the spatial quality of pitch (from Helmholtz onwards) and make an important distinction between the peculiarly musical and the auditory types of space. The auditory space is the external, physical place where sounds sources are perceived to be located. By contrast, sound space is a phenomenal and non-empirical space to which listeners (might) refer to when describing sounds or music they listen to. It might rely on musical or acoustic features of sounds, such as dynamics and pitch, but it resists to the definition in mere psychoacoustic terms. Second, they highlight the crucial role of the notion of movement in grounding the conceptualization of auditory perception on a spatial basis. Third, they necessarily fail when trying to provide the notion of musical space of a solid and rigorous definition, so that the nature of such musical spaces remains controversial if not unclear (Kania, 2015).

To properly understand the notion of musical space, therefore, we must not restrict *a priori* the notion of space to the commonsensical idea of physical space, i.e., a geometrical structure in which objects exist as material entities[1]. On the contrary, exploiting the rich variety of different conceptualizations of space in different fields, such as physics, mathematics, and geometry (see Jammer, 1993), we can admit the existence of a peculiar sonic, or musical, space as a non-empirical and phenomenologically experienced space whose perceptual features can be, at least to some extent, described. This goes back to the n-dimensional

[1] See Mattens (2018) for a phenomenological analysis of the problem of spatiality and the multiplicity of constructs of space.

manifold mentioned by Helmholtz (see also Deutsch, 1984). However, the question of dimensionality seemingly remains open, namely regarding the number of dimensions that are really needed to describe the full character of musical "space".

4. Spatial compositions

Composers of the second half of the twentieth century went deeper into the spatial nature of sound. Interesting insights come from the works of Edgar Varèse and of the so-called spectral composers. In Varèse's *Ionisation*, different elements of sound are projected into a dynamic acoustic space which forms the auditory environment of the composition. For Varèse, music is intimately 'spatial', and it is so due to the movement of resounding masses varying in intensity and density. Similar ideas will form the basis of spectral music, in which sonorous masses dynamically evolve in time and move in space.

The concept of space plays a fundamental role in both the theorizations and musical works of the Greek composer Iannis Xenakis. Solomos (2015) presents the notion of space in Xenakis's thought and artistic production according to four levels (see also Harley, 2004). At the philosophical level, Xenakis occasionally considered time as an epiphenomenon, and space as a more fundamental reality (1988). At the compositional level, geometrical space is used as a tool to inspire and create musical compositions, in which the sonorous object takes on spatial forms. The formalization of music might take advantage of graphic objects (e.g., dots and lines). At the physical level, the spatialization of sound becomes a fundamental factor in musical composition. The fourth level deals with a more concrete notion of space, related to the physical environment and to the concept of "place", which allows us to understand the need for space in terms of aesthetics as a way to rediscover the multiplicity of meaning.

The understanding of sound as a concrete, and in some cases even graphic, spatialized entity demonstrates a general tendency towards imagining space in terms of its subsequent implementation and filling. Spatial forms in many works of Xenakis are a prototype of how order is introduced into primary chaos. Music, conceived of as sonorous materials developing in time, occurs in a space and creates the phenomenon of internal sound space. In *Alax* (1985), for instance, three identical ensembles form an equilateral triangle around the audience. This arrangement enhances, through spatial separation, the perception of simultaneous and independent layers of sound. It also allows the creation of spatial canons, echoes, and overlapping sonorities as well as providing effective perceptual cues for sound movement. The balance of sound-masses and the trajectories of sound movements will seem different to the audience according to their different spatial locations. With respect to Nineteenth-century music, in which space is mostly an accidental factor of music performance, several Twentieth-century composers conceive the musical work primarily as spatial, with space being a key variable of its representation and fruition.

One of the most influential compositions in the history of spatialized music is *Gruppen fur drei Orchester* (1955-1957) by Karlheinz Stockhausen. About a hundred musicians are split into three orchestras (each with its own conductor) arranged in an arc position on platforms semi-surrounding the audience in a large hall. Sound masses are produced and moved from one orchestra to the other, often changing their acoustic profile. The illusion of movement is constructed by a temporal overlapping and dynamic shaping of the sounds. The chords swell dynamically from one orchestral group to the other, creating an impression of the continuation of the sound essential for the perception of its movement, trying to provide the audience, placed in the center, with the impression of a rotating sound.

The above examples illustrate how the conceptualizations of music in spatial terms have impacted compositional practices, with most occurrences seemingly rely upon a quite literal interpretation of the relationship between music and space, by exploiting space as a non-auditory compositional feature for music[2]. However, as we have seen above, the same relationship has been occasionally interpreted in conceptual or metaphorical terms, thus suggesting that the link between music and space is mediated by a cognitive factor. In the next section, I examine some of the issues emerging from both the perceptual and cognitive interpretation of the music-space relationship.

5. Issues emerging from the analogy between music and space

Some spatial conceptions of music reviewed here assume the possibility of establishing a similarity between visual and music percepts (e.g., Helmholtz and Wellek, see also Koffka, 1935; Mach, 1894, 1906; von Ehrenfels, 1890). The explanations of such analogy have often evoked the mediating role of movement, which can intuitively be attributed to pitches in the melodic space as well as to objects in the physical space (Di Stefano, 2022; Larson, 2012; Hubbard, 2017). At the same time, however, literature in the field of crossmodal perception has pointed out the issues that emerge when similarity relationships are evoked to link stimuli across different senses, such as vision and audition (Di Stefano & Spence, 2023; Spence & Di Stefano, 2023). While it is commonly accepted that similarity might exist between pairs of stimuli presented within the same sensory modality, conceiving sensory similarity between pairs of stimuli presented in different sensory modalities would appear to be a much more controversial

[2] A further occurrence of the music-space relationship deals with the disposition of musicians in different orchestras, which impacts the way sound is generated and therefore experienced by the listeners.

topic, with the question of how to explain similarity in the crossmodal domain attracting a diverse range of theoretical responses (Spence, 2022, for perceptual similarity in the chemical senses).

On the one side, Lawrence Marks has suggested that the different senses work in close communication and have to be considered as a potentially unified, or unifying, processing system (Marks, 1978). Commenting on some of his own findings on audiovisual associations, Marks writes: "In a cross-modality matching task, for example, virtually all subjects will set higher sound frequencies to match greater visual intensities (Marks, 1974, 1978), thereby revealing a universal appreciation of similarity between the dimension of pitch on the one hand and that of brightness on the other" (Marks, 1989, p. 58). Moreover, Marks (1987) observed that crossmodal associations might be due to some kind of communality between parallel stages in processing attributes of different sensory modalities; such communality may not be explained in terms of phenomenological properties of the stimulus, but rather in terms of stimulus processing after sensory apprehension.

On the other side, other commentators have refuted the idea of establishing similarity on sensory basis, given that the senses process and convey different kinds of information. Helmholtz was himself skeptical in this regard, observing that: "The distinctions among sensations which belong to different modalities, such as the differences among blue, warm, sweet, and high-pitched, are so fundamental as to exclude any possible transition from one modality to another and any relationship of greater or less similarity. […] Comparisons are possible only within each modality" (Helmholtz, 1878/1971, p. 77). Similar statements seemingly exclude that the relationship between space and music can be mediated by visually perceived aspect, such as movement. Moreover, even assuming that phenomenal quality can be shared by objects across different modalities, objections have been raised that have pointed to the fact that, given that any two objects share at least one phenomenal quality, similarity would simply be a universal relation and therefore similarity claims are uninformative (Goodman, 1972).

Given the difficulties in establishing perceptual similarity on a purely sensory/psychophysical basis, several psychologists have suggested an alternative, cognitively imbued view of similarity, which could seemingly apply to the case of space-music analogy. According to this view, the analogy between space and music would be established more on conceptual criteria, such as the geometrical rules that allow for the representation of tridimensional realities as bidimensional (e.g., the Mercator projection). Cross-species evidence suggested that the existence of a general analogy factor might be related to the way in which evolutionary forces have shaped human brains to detect sensory regularities in environmental stimuli (Ravignani & Sonnweber, 2017). Together with the findings that have been reported in infants (Lewkowicz & Turkewitz, 1980), the evidence suggests that basic cross-modal abstraction capacities may involve evolutionary, low-level sensory mechanisms.

The problem of the cognitive/perceptual nature of similarity across the senses might be also framed from Alistair's (2013) distinction between property-based and isomorphism-based accounts of similarity. The property-based account treats two objects as similar when they share phenomenal qualities. The isomorphism-based account treats two objects as similar if a mapping between their component parts exists. While the idea of property-based similarity appears to be intuitively and straightforwardly applied in the unisensory domain, in the crossmodal domain it seems that isomorphism-based accounts have more explanatory power. As suggested by Di Stefano & Spence (2023), referring to the concept of perceptual salience can also explain hierarchies in the way crossmodal similarities are established by the perceiver, with the sensory ones being used first, if available, while the most cognitive ones are triggered only when perceptual similarities are unavailable.

6. Conclusions

The exploration of the temporal and spatial dimensions in sound perception has unveiled intricate relationships that go beyond the mere auditory experience. The temporal nature of sound perception reveals the human ability to parse a continuous auditory flow into distinct perceptual units thanks to Gestalt principles. The psycho-musicological models inspired by the temporal dimension underscore how sounds fulfil expectations within the perceptual flow, emphasizing the key role of time in the conceptualization of sounds by both composers and listeners.

Temporal exploration introduces a paradox challenging the traditional "now," expanding present duration and blurring past and future in melody perception. Transitioning to spatial concepts of sound, the emergence of a unique musical space is discussed by several philosophers and musicologists, from Helmholtz to contemporary scholars. The notion of musical space, whether as a conceptual structure or a non-empirical space, remains elusive, prompting an interdisciplinary exploration that draws from physics, mathematics, and geometry. As the discussion extends to 20th-century music, it becomes evident how several composers utilized spatial elements as compositional features.

However, the analogy between music and space faces challenges in establishing perceptual similarity solely on sensory information. The debate between property-based and isomorphism-based accounts of similarity highlights the role of cognitive factors in establishing analogies between space and music.

In conclusion, the temporal and spatial dimensions of sound perception offer a rich tapestry of exploration, intertwining psychology, philosophy, and musicology.

The dynamic interplay between past and present in temporal perception and the elusive nature of musical space challenge traditional notions, calling for a multidimensional investigation of how we experience and conceptualize sound.

Summary

Music has been primarily conceived as a temporal art. However, over the last two centuries or so, researchers across different disciplines including musicology, psychology, and philosophy, have been intrigued by the spatial nature of music and sounds, using spatial concepts to define music. This paper aims to demonstrate that an understanding of music perception from a temporal perspective inherently implies a certain spatial dimension. To do this, first, I briefly examine some key arguments that lead to conceiving sound perception in temporal terms. At the same time, I highlight some of the limitations of a purely temporal account of sound perception which necessitate the incorporation of spatial considerations into the conceptualization of sound perception. Consequently, I move on to consider prominent spatial accounts of musical sounds that have been elaborated by psychologists, musicologists, and music composers. In conclusion, I discuss some of the challenges arising from the analogy between music and space, whether conceived in perceptual or cognitive terms.
Keywords: Auditory perception; Phenomenology; Musical space; Spatialization of sound.

Zusammenfassung

Musik wurde in erster Linie als eine zeitliche Kunst betrachtet. In den letzten etwa zwei Jahrhunderten haben jedoch Forscher aus verschiedenen Disziplinen, darunter Musikwissenschaft, Psychologie und Philosophie, das Interesse an der räumlichen Natur von Musik und Klängen geweckt und räumliche Konzepte verwendet, um Musik zu definieren. Diese Arbeit zielt darauf ab, zu zeigen, dass eine Betrachtung der Musikwahrnehmung aus einer zeitlichen Perspektive notwendig eine gewisse räumliche Dimension impliziert. Um dies zu erreichen, untersuche ich zunächst kurz einige Schlüsselargumente, die dazu führen, dass die Klangwahrnehmung in zeitlichen Begriffen betrachtet wird. Gleichzeitig hebe ich die Grenzen einer rein zeitlichen Darstellung der Klangwahrnehmung hervor, die die Einbeziehung räumlicher Überlegungen in die Konzeptualisierung der Klangwahrnehmung notwendig macht. Infolgedessen werfe ich einen genaueren Blick auf bedeutende räumliche Darstellungen musikalischer Klänge, die von Psychologen, Musikwissenschaftlern und Musikkomponisten ausgearbeitet wurden. Abschließend gehe ich auf einige Herausforderungen ein, die aus der Analogie zwischen Musik und Raum entstehen, sei es in perzeptiven oder kognitiven Begriffen.
Keywords: Auditive Wahrnehmung; Phänomenologie; Musikalischer Raum; Verräumlichung von Klang.

References

Alistair, M. C. I. (2013). Objective similarity and mental representation. *Australasian Journal of Philosophy*, *91*(4), 683-704.

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, *14*(10), 693-707.

Bregman., A. S. (1990). *Auditory scene analysis: the perceptual organization of sound*. The MIT Press, Cambridge, MA.

Casati, R., Dokic, J., 1994: *La philosophie du son*, Chambon, Paris.
Casati, R., Dokic, J., Di Bona E., 2020: Sounds, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/win2020/entries/sounds/
Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, *25*(5), 975-979.
Clarke, D. (2011). Music, phenomenology, time consciousness: meditations after Husserl. In E. Clarke, D. Clarke (Eds.), *Music and consciousness: Philosophical, psychological, and cultural perspectives*, Oxford University Press (pp. 1-28).
Clifton, T., 1983: *Music as heard. A study in applied phenomenology*, Yale University Press, New Haven.
Di Stefano, N. (2022). The spatiality of sounds. From sound-source localization to musical spaces. *Aisthesis. Pratiche, linguaggi e saperi dell'estetico*, *15*(1), 173-185.
Di Stefano, N. (2023). Musical Emotions and Timbre: from Expressiveness to Atmospheres. *Philosophia*, 1-13. https://doi.org/10.1007/s11406-023-00700-6
Di Stefano, N., & Spence, C. (2023). Perceptual similarity: Insights from crossmodal correspondences. *Review of Philosophy and Psychology*, 1-30.
Deutsch, D. (1984). Musical space, *Advances in Psychology*, (19), 253-287.
Drake, C., & Bertrand, D. (2003). The quest for universals in temporal processing in music. In R. J. Zatorre & I. Peretz (Eds.), *The cognitive neuroscience of music* (pp. 21-31). Oxford: Oxford University Press.
Godøy, R. I. (2021). Perceiving sound objects in the musique concrète. *Frontiers in Psychology*, 12, 1702.
Goodman, N. (1972). Seven strictures on similarity. In *Problems and projects* (pp. 437-446). Indianapolis, IN: Bobs-Merril.
Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, *5*(11), 887-892.
Harley, J. (2004). *Xenakis. His life in music*. Routledge.
Harley, M.A., 1994: *Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations*, Ph. D. dissertation, McGill University.
Hegel, G.W.F., 1835: *Aesthetics. Lectures on fine art*, trans. by T.M. Knox (2 Vols.). Oxford University Press, Oxford (UK), 1975.
Hubbard, T.L. (2017). Momentum in music: Musical succession as physical motion. *Psychomusicology: Music, Mind, and Brain*, 27(1), 14-30.
Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
Husserl, E., 1928: *On the phenomenology of the consciousness of internal time*, Trans. By J. Barnett Brough, Kluwer, Dordrecht, 1991.
Jammer, M., 1993: *Concepts of space: the history of theories of space in physics*, Dover edition, New York.
Kania, A., 2015: An imaginative theory of musical space and movement, *British Journal of Aesthetics*, *55*(2), 157-172.
Koffka K. (1935). *Principles of Gestalt psychology*. New York.
Kubovy, M. (1988). Should we resist the seductiveness of the space: time:: vision: audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 318-320.
Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*(1-2), 97-126.
Kurth, E. (1931). *Musikpsychologie*, Hesse, Berlin.
Larson, S. (2012). *Musical forces: Motion, metaphor, and meaning in music*, Indiana University Press, Bloomington.
Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Developmental Psychology*, *16*(6), 597-607. https://doi.org/10.1037/0012-1649.16.6.597.
Lippman, E.A. (1952). *Music and space. A study in the philosophy of music*. Ph.D. dissertation, Columbia University, New York.
Macedo, F. (2015). Space as metaphor: the use of spatial metaphors in music and music writing. *Signata. Annales des sémiotiques/Annals of Semiotics* (6), 215-230.
Mach, E. (1959). *The Analysis of Sensations and the Relation of the Physical to the Psychical*. (C.M. Williams, transl.). New York: Dover. (Originally published in Germany, 1906).
Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, *22*(4), 663-714.
Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American Journal of Psychology*, *87*, 173-188.
Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. New York, NY: Academic Press.
Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 384-394.

Marks, L. E. (1989). For hedgehogs and foxes: Individual differences in the perception of cross-modal similarity. In G. Ljunggren & S. Dornic (Eds.), *Psychophysics in action* (pp. 55-65). Berlin: Springer Verlag.

Mattens, F. (2018). *From the Origin of Spatiality to a Variety of Spaces*. In Zahavi, D. (Ed.), *The Oxford handbook of the history of phenomenology*, Oxford University Press, New York, pp. 558-578.

Matthen, M. (2010). On the diversity of auditory objects, *Review of Philosophy and Psychology*, 1(1), 63-89.

McDermott, V. (1972). A conceptual musical space, *The Journal of Aesthetics and Art Criticism*, 30(4), 489-494.

Morgan, R.P. (1980). Musical time/musical space, *Critical inquiry*, 6(3), 527-538.

Moser, H.J. (1953). *Musikästhetik*, De Gruyter, Berlin.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.

Nudds, M. (2009). *Sounds and Space*, in Nudds, M., O'Callaghan, C. (Eds.), *Sounds and Perception: New Philosophical Essays*, Oxford University Press, Oxford, pp. 69-96.

O'Callaghan, C. (2009). Seeing what you hear: Cross-modal illusions and perception, *Philosophical Issues*, 18, 316-338.

Pesic, P. (ed.) (2007). *Beyond Geometry: Classic Papers from Riemann to Einstein*, New York: Dover Publications.

Pesic, P. (2013). *Helmholtz, Riemann, and the Sirens: Sound, Color, and the "Problem of Space"*, Physics in Perspective, 15, 256-294.

Ravignani, A., & Sonnweber, R. (2017). Chimpanzees process structural isomorphisms across sensory modalities. *Cognition*, *161*, 74-79.

Riedel, F. (2019). *"The Atmospheres of Tones": Notions of Atmosphere in Music Scholarship Between 1840 and 1930*. In Griffero T., Tedeschini, M. (Eds.) *Atmosphere and Aesthetics*, Palgrave Macmillan, Cham, pp. 293-312.

Schaeffer, P. (1966). *Traité des Objets Musicaux*, Seuil, Paris.

Schoenberg, A. (1950). *Style and Idea*, Philosophical library, New York.

Solomos, M. (2015). *The Complexity of Xenakis's Notion of Space*. In M. Brech, R. Paland (éd.), *Komposition für hörbaren Raum. Die frühe elektroakustische Musik und ihre Kontexte* [Compositions for Audible Space. The Early Electroacoustic Music and its Contexts, Bilefeld], transcript Verlag, p. 323-337.

Spence, C. (2022). Searching for perceptual similarity within, and between, the (chemical) senses. *i-Perception*, *13*(5):20416695221124154.

Spence, C., & Di Stefano, N. (2022). Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and sound, *i-Perception*, 13, 1-42.

Spence, C., & Di Stefano, N. (2023). Sensory translation between audition and vision. *Psychonomic Bulletin & Review*, 1-28.

Stravinsky, I. (1947). *Poetics of music in the form of six lessons*, Harvard University Press, Cambridge.

Tenney, J., & Polansky, L. (1980). Temporal gestalt perception in music, *Journal of Music Theory*, 24(2), 205-241.

Von Ehrenfels, C. (1890). Uber Gestaltqualitaten. *Viertel jahrschrift fur Wissenschaftliche Philosophie*, 14, 249-292.

Von Helmholtz, H. L. (1878/1971). *Treatise on physiological optics* (Vol. II). New York: Dover Publications.

Von Helmholtz, H. (1954). *On the sensations of tone*, translated by A.J. Ellis, New York: Dover Publications.

Warren, R. M. (2013). *Auditory perception: A new synthesis*. Elsevier.

Wellek, A. (1963). *Der Raum in der Musik*. Appendix *to Musikpsychologie und Musiklisthetik. Grundriss der Systematischen Musikwissenschaft*, Akademische Verlagsgesellschaft, Frankfurt am Main.

Xenakis, I. (1988). *Sur le temps*, in Xenakis, I., *Kéleütha*, ed. by Alain Galliari, Paris 1994.

Zbikowski, L. M. (2002). *Conceptualizing music: Cognitive structure, theory, and analysis*, Oxford University Press, Oxford.

Zuckerkandl, V. (1956). *Sound and symbol: music and the external world*, Pantheon Books, New York.

Nicola Di Stefano is Researcher at the National Research Council of Italy. His research activity focuses on topics at the intersection between aesthetics, music, and philosophy of perception. He is adjunct professor at the University of Arkansas Rome Center and the Campus Bio-Medico University of Rome. He is currently involved in several EU-funded projects on topics related to robotics applications to music (CONBOTS), human augmentation (NIMA), and neurotechnologies (B-CRATOS).

Address: Institute of Cognitive Sciences and Technologies (ISTC) National Research Council of Italy (CNR) Via San Martino della Battaglia 44, 00185, Rome, Italy

E-mail: nicola.distefano@istc.cnr.it

ORCID: 0000-0002-9286-0395