

Memory as the Aspatial Domain for the Perception of Certain Genres of Contemporary Art Music

Music & Science
Volume 4: 1–18

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/2059204321997658

journals.sagepub.com/home/mns

Riccardo Wanke¹ and Vincenzo Santarcangelo²

Abstract

This paper enquires into the nature of the connections between memory and certain genres of contemporary art music whose unique features rely particularly on our early mnemonic processes. Specific sound configurations of this music are often associated, during listening, with visual and tactile sensorial qualities and with abstract geometries. They are perceived fundamentally as the results of acoustic-physical forces and energies and are organized according to Gestalt and kinesthetic principles. This kind of music calls for a specific listening attitude, which we define as the vertical stance, and seems particularly apt to respond to mechanisms of the working memory where echoic, short- and long-term memories assume a central role. In this vertical stance, memory is involved in the mental construction (segregation, storage, and prediction) of the Gestalt configurations of this music within a perceptual domain that crucially has no spatial connection to the external world. In tying in neurophysiological and psychological research with musicological theories, we discuss the perceptual approach to these music practices in the light of the philosophical concept of the 'No-Space world' as conceived by the philosopher Peter Strawson. We propose that – under certain conditions – memory may be the realm of the purely spectro-temporal features of music. The sound configurations of this music in particular are part of an internal-external perceptual framework, being decoded in the conceptual space of perception and able to elicit high-order recollections typical of an embodied engagement with the external world.

Keywords

Memory, contemporary art music, auditory perception, Gestaltic formation, auditory scene analysis

Submission date: 02 February 2021; Acceptance date: 02 February 2021

Introduction

The rich connections between music and memory assume a particular significance when we start to consider contemporary art music¹ (from now on CAM) and its listening practices. How is contemporary art music related to mnemonic abilities? When we start to consider CAM, the discourse around memory relating to the more common genres of Western classical, folk, or world music does not seem to fit. The idea behind this paper is that the specific features of pieces belonging to certain genres of CAM, such as post-spectralism, minimalism, electroacoustic music, and various offshoots of electronica, may engage listeners in a different way as compared to other forms of music.

On the one hand, both the variety and the limited diffusion and consumption of CAM hinder the consolidation of a widespread repertoire that would assist in the identification

and retention of various features of the music by a listener. On the other hand, it is evident that some 20th century music, due to its unconventional character and inventive nature, inherently demands a specific outlook capable of grasping the distinctive facets of these music practices (Solomos, 2019). Traditional '[e]xamples of ways to make pieces difficult to represent [mentally] would be: building a piece with virtually no repetition [...]; building a piece so that there

¹ CESEM – Centre for the Study of the Sociology and Aesthetics of Music, FCSH, NOVA University, Lisbon, Portugal

² Polytechnic of Turin, Italy

Corresponding author:

Riccardo Wanke, CESEM – Centre for the Study of the Sociology and Aesthetics of Music, FCSH, NOVA University, Lisbon 1069061, Portugal.
Email: rwanke@fcsh.unl.pt



appears to be only one repeating schema with variations [...] building a piece out of indefinite and uncategorizable events such as indistinct pitches, slides, noises, etc.’ (Snyder, 2001, p. 101). These are just some of the characteristics of aleatoric music, free improvisation, minimalism and drone music, electroacoustic, noise and contemporary classical music practices. The absence of tonal constructions, traditional narrative and temporalities, and the presence instead of real-world sounds, noises and distorted sounds, acoustic and electronic sounds, all potentially within non-linear temporal constructions, hampers conventional descriptions and challenges the research to find new models able to explain the mechanisms behind our apprehension of this music.

While the appreciation of this kind of music still largely follows general mechanisms that are dependent on familiarity and background (Bourdieu, 1979; Grebosz-Haring & Weichbold, 2020; Menger, 2001), our aim here is to discuss how these genres of music escape conventional forms of memorization and to examine the impact of this music on the most essential areas of our perception, exploring new perceptual pathways that these genres engage.

In this paper, we consider certain genres within CAM that show an explicit interest in the intrinsic properties of sound, rather than in any organized and planned construction. The focus of these genres revolves around the creation (and the experience of) sonic textures and masses in motion, instead of a dialectic relation between tones (i.e., melodies). These pieces distinctly show a passage from sound seen as a *constituted* element of musical phrases with a given function to sound as *constituent* element of music, autonomized in such a way that we may define it not through its usage but through its nature (Bonnet, 2012, p. 130). Listening to this music, which spans electroacoustic music, mixed-source works, post-spectralism, minimalism, glitch-electronica, drone music, microsound, sound art, and dubstep, often leads to associations with visual and tactile sensorial qualities and to abstract geometries (e.g., surfaces, planes, lines, and points) (Godøy, 1997; Wanke, 2019), due essentially to the sonic characteristics of these musical genres (Kendall, 2014; Kiss, 2004). This paper examines how memory is involved and operates when we listen to these specific music practices.

The key aspect which differentiates these styles of music from others is that, apart from being grounded in the intrinsic properties of sound, this music displays essential sonic forms, that we define here as *sound configurations*, arranged and constructed following the principles of Gestalt (Santarcangelo & Wanke, 2020). We aim to show how the nature of these configurations is tied to the early stage of perception and is strictly connected to the functioning of working memory (WM) (Joseph et al., 2015). By describing our primary response to sonic information which results from the combination of both the echoic and short-term memory (STM) and auditory Gestalten formation (Koelsch & Siebel, 2005, p. 579), we focus on how these configurations prompt a particular response from our

memory, favoring a specific perceptual stance. They impact physically on our echoic memory, fortify STM, and control the formation of long-term memories (LTMs).

If this music is based on Gestalt features, what do we retain in memory when we remember an *organized sequence of sounds*²? We suggest that these sound configurations are part of what we define as an internal-external perceptual framework, being recognized and processed by Gestalt principles of perception, stored by WM, and acting as bridging elements for an embodied cognitive response. The importance lies in the fact that listening to this music goes beyond a cognitive abstraction (i.e., extraction of invariants) for the formation of corresponding mental representations in a listener’s mind. Instead, by matching efficaciously the auditory Gestalten processing at the early stages of perception, these sound configurations detach themselves from their source of production, allowing for a morphodynamic analogy between the external acoustic features and the internal mental images (Reybrouck, 1997).

In giving a full description of this process, we attempt to trace a connection between music, cognitive science and philosophy of perception. Our argument relies on the concept of the ‘No-Space world’, the idea developed as a thought experiment by Peter Strawson (1959). The auditory system detects and identifies sources, but when we approach music perceptually, and in particular this kind of music, its sounds appear to detach themselves from any actual origin and come to occupy a precise and distinct perceptual domain. The experience of this music entails a ‘vertical’ perceptual state of consciousness (Kozak, 2020; Kramer, 1988) in which primal impressions and the retention and protention of sound configurations are connected within a perceptual domain. This domain, realized within the functioning of WM, corresponds to an *aspatial* domain of perception where the image schemata of this music recall experiential associations reaching out towards an embodied cognition.

Our aim with this paper is to propose an alternative perspective that is able to explain our perceptual approach to a series of music practices which elaborate unconventional sonic forms. The sonic stream that reifies itself in auditory memory is actually an aspatial domain where only pure sounds may exist.

Sonic Characteristics

Certain genres of CAM are intrinsically based on essential sonic forms, rather than any complex and functional arrangement. This tendency has emerged in different musical currents over the past century and may be outlined as ‘the process of emancipation of sound within the music practices of the 20th century’ (Solomos, 2019). We focus on those music genres which develop a series of practices – such as the restoration of a temporal continuity, the integration of different sound sources, and the extended exploration of sonic spectrum – as cardinal principles,

including spectralism, minimalism, avant-garde music, electroacoustic compositions and certain offshoots of experimental improvisations, and electronica.

If we consider pieces by figures such Alvin Lucier, Georg Friedrich Haas, Jacob Ullman, Éliane Radigue, Ryoji Ikeda, Pan Sonic, Jürg Frey, or Richard Chartier, we encounter a set of works that spill over the limits of our perception. Retention and protention – following Husserl’s terminology – of the sonic flow in these compositions are *intentionally* put to the test. These compositions use material that in fact exceeds the definite form of a musical (or sound) object: this material does not possess clear edges or boundaries (being potentially confused with environmental sound or creating acoustic challenges in perception), it puzzles our perceptual segregation by shifting the intensity of similar aural stimuli or by mixing different sounds into a coherent *perceptual chain* (i.e., families of morphologically similar sounds), challenging our expectation of future musical events. Some of these works use such primordial structures that they minimize our predictive constructions. Jürg Frey’s string quartets (e.g., second and third) are exemplary in this sense. In composing a repetitive and monolithic sequence of suspended sounds, Frey reduces our processing of sonic streams to its essentials, favoring a static apprehension of sounds. The listener is immersed in a consistent state where dynamic and ADSR (attack, decay, sustain, release) features of sound are uniform, allowing the listener to take these elements for granted and thus be able to focus on spectral characteristics.

One of our previous study (Wanke, 2015) illustrates how authors belonging to these genres center their work on a series of specific sonic morphologies and evolutions. Morphological aspects span a large variety of sound types: the use of microtonality and sonic layers; the construction of sonic masses; the creation of systematic continuous pitch movements (i.e., *glissandi*) and repetition. These characteristics are embedded within musical evolutions such as metric structures and the opposition of different elements according to certain sonic dimensions (e.g., dynamic or timbric). Composers and performers within these genres employ a ‘plastic’ use of sounds based on their various identities – whether according to their sonic properties, spatial character, their role, intensity, or presence. In this sense, sounds appear and emerge possessing a sort of tactile and sculptural presence within a depicted sonic domain of perception as they occupy a particular spaciousness. These artists use a limited set of elements that are part of a unique conception of sonic flow and they focus on the presentation of aural elements minimizing their articulation and mutual interplay. These pieces exhibit a reduced dialectic between musical elements that behave more as structural constituents of a complex sound rather than as part of a narrative exchange. Therefore, given the potential variety of sound types, this music reflects an idea of convergence of elements and behaviors. This idea is more evident in those pieces which include continuous musical episodes made

up of sustained and overlapping tones, building a kind of material thickening, a layering up in perception. When discrete sounds are employed, pieces (or episodes) proceed by developing families of sounds in such a way as to create a natural internal coherence in which each aspect is perceived as part of an integrated form.

Most essentially, this kind of music is based on ineffable and overwhelming sonic dimensions accessed through sonic phenomena at the threshold of hearing (microsound and glitch-electronica), transient impressions (spectralism), distorted sonic aggregates (noise, electroacoustic music), or even specifically constructed illusory effects (minimalism).

Perceptual Grammar

We argue that, due to the unfamiliar constructions present in these genres of CAM, this type of music engages our memory more easily at a phenomenological and sensorial level rather than through establishing long-term associations. In referring to this kind of music, Snyder talks about a ‘memory sabotage’: a music that tries to exist outside of memory by flouting the structural principles of more conventional music, that is ‘by constructing patterns whose distribution of information makes them difficult to process, or [...] by using events or silences whose time lengths exceed the limits of short-term memory’ (2001, p. 235). It may be difficult, for instance, to remember exactly long musical passages from pieces such as *String Quartet n°4* (1964) by Scelsi, *Trilogie de la Mort* (1993) by Radigue, *Lontano* (1967) by Ligeti, or +. (1996) by Ikeda. These pieces possess an ambiguous evolution and frequently consist of a continuous ever-changing sonic stream where liminal and ephemeral sonic effects pop up without the listener being fully aware of them. Solomos, for instance, describes listening to Scelsi’s *String Quartet n° 4* as a kind of invention: ‘not only do I hear different things each time I listen, in addition, it is always difficult for me to know whether what I hear (or think I hear) is “there” (in the score or in the recording) or whether my mind is inventing it’ (Solomos, 2019, p. 126). It is more likely that one would have a precise recollection of colors and chord combinations than of extended temporal or formal structures. There are still, of course, rhythmic patterns, distinct evolutions and sonic gestures which may leave vivid memories, but the restricted dialectic and the overall continuity and unicity of sound both favor the memorization of brief episodes over larger structures and the recollection of sonic situations of sound elements arranged within a spectro-temporal range rather than the equivalent of a precise sequence of notes and melodies.

The characteristics of this music account for a field of sound configurations covering both structural (e.g., extended spectral exploration, microtonality, repeated *glissandi*, sonic masses) and unfolding characteristics (e.g., rhythmic and reiterated figures, sonic contrasts), their effects (e.g., hypnotic, cumulative) as well as their arrangements (e.g., non-functional, sculptural, limited variety,

Table 1. Sonic characteristics and typologies (Wanke, 2015).

Types	Characteristics
Morphologies	<ul style="list-style-type: none"> • The use of an expanded spectrum • Interactions between neighboring frequencies (e.g., microtonality, binaural beats) • Systematic frequency-based glides (<i>glissando</i>) • Static masses of sound (e.g., sustained tone-based aggregates, drones)
Evolutions	<ul style="list-style-type: none"> • Metric structures (e.g., rhythm, repetitive clusters) • Dynamic and timbric contrasts
Effects	<ul style="list-style-type: none"> • Hypnotic reiterations • Auditory challenges at macro/micro sound scales (e.g., distortions, illusory effects, psychoacoustic phenomena)
Arrangements	<ul style="list-style-type: none"> • A plastic and sculptural arrangement of sound • Restricted number of elements conceived globally • Limited dialectic among elements

extended temporal and spectral presentations) (see Table 1). Taken as a whole, these sound configurations relate to the so-called intrinsic properties of sound. This is the portion of the acoustic stimulus where sound expresses its inner content, its ‘colours, light effects, tactile sensations, all properties of a physical matter’³, and where it possesses a quasi-objective presence and material nature: it is the place of the palpable qualities of sound, its volume, its density and grain (Bériachvili, 2010, p. 32). In this portion of the acoustic stimulus, we experience a direct appraisal of intrinsic sonic features and a reduced ability in spatial localization and origin identification. The use of complex tone aggregates, the unconventional decay of inharmonic series, the presence of noises and distortions puzzle in fact phase discernment and sound localization (i.e., bases of binaural hearing). The use of microtonal interactions, in particular, leads to suspended non-linear patterns that ‘include those parametric values [which] do not change over time, but merely repeat; fluctuate between fixed values in no particular order’ (Snyder, 2001, p. 65). The acoustic phenomenon of binaural beats – a beat originated by an interference between two sounds of slightly different frequencies and frequently used in spectralism, and materialist minimalism (Corbett, 2000) – is an acoustic illusion which has a similar basis (i.e., phase difference) to binaural hearing – that is, our ability to locate sound ‘out there’. Precisely because of this common ground, binaural beats challenge our capacity to recognize a sound’s position in the outer world and can generate the sensation that sounds contributing to this interference pattern are located somewhere in the head (Oster, 1973). Moreover, it is worth resolving some of the ambiguity in the use of the term ‘space’ in relation to sound. Among the large variety of possible meanings⁴, here we use this term to refer to the outer physical space. In order to avoid confusion, we denote a general three-dimensional

extent, such as that of the frequency/time/intensity set of parameters, as a ‘domain’. This distinction may help the reader in grasping the key difference between the domain of a figure/background sonic construction versus, for example, the spatial direction of a sound of a car passing by ‘out there’.

Due to their nature, these sound configurations are often used in combinations (e.g., glitch-electronic music usually exhibits repetitive clusters within rhythmic frameworks, while the repetition of musical units in longer pieces of contemporary classical music can at times be associated with non-rhythmic hypnotic reiterations or more complex structures (Wanke, 2015)) which are typical of interdependencies and hierarchies. In this sense, they form a sort of *perceptual grammar* which incorporates internal principles and synergies. The concept of perceptual grammar parallels Kanizsa’s visual grammar (1979), insofar as this perceptual grammar embodies Gestalt principles of proximity and similarity and embraces a series of essential (identity, difference, motion, and stasis), structured (ascending/descending, figure/background), and dynamic or kinesthetic (tension, linearity, amplitude, projection) patterns (Cruys & Wagemanns, 2011; Sheets-Johnstone, 1999; Tenney, 1988). Indeed, the very first encounter with this music reproduces forms shaped by spectro-temporal boundaries (Kubovy & Van Valkenburg, 2001), such as figure/ground segregation, complex multiple plane arrangements, clashes between pulses vs. sustained stimuli, time-based frames eliciting causal chains in perception, and the unfolding of contrasting profiles. The essential feature that unites these forms is their abstract nature: they do not resemble anything in the world but instead, as will be made clear later on, possess a fundamental link to the world in terms of potential.

A Memory Sabotage or a New Engagement

Characteristics of this perceptual grammar – which traditional studies on music perception attribute to a kind of ‘memory sabotage’ (Snyder, 2001) – are the essential elements of a series of music practices today. This music does not search for unexpected resolutions of conventional schemes of expectation (Huron, 2006; Raffman, 2003), but rather points to another perceptual appraisal, a *new engagement*, that Kramer described as a ‘vertical perspective’ (1988) of listening and to which Meyer alludes as the ‘simultaneous deviation’ (1956). If ‘musical groupings [of more conventional genres] are often created by composers so as to fit within the limitations of STM, thus giving their music a clearer and more memorable structure’ (Snyder, 2001, p. 37), this music intentionally pushes listeners to find more uncommon perceptual approaches. This shift of perspective is evident if we consider composers such Xenakis, Grisey, or Lucier or performers such Autechre and Jacob Kirkegaard. These figures incorporate psychoacoustic phenomena in their pieces (e.g., Kirkegaard’s

Labyrinthitis (2008) exploits the phenomenon of the evoked otoacoustic emission) and the study of neural networks to design algorithms for composition (e.g., Xenakis, Vaggione, Autechre) (Kiss, 2004). The general idea behind the work of these authors is to stimulate the audience to find other ways to listen which may include particular listening attitudes (Grisey, López, Oliveros) or the resonance of physical space (Xenakis, Lucier) and therefore expand toward immersive experiences (electroacoustic, sound art and soundscape). Many composers working in this area explicitly reject the idea of music expressing a meaning or representing something (e.g., an emotion) in favor of an idea of music as stimulating an aural experience that is engaged with by each listener in their own way (López, 2004; Oliveros, 2005).

The focus here is on sound and its impact on our perception. Grisey himself explains that '[t]he more we expand our auditory acuity to perceive the microphonic world, the more we draw in our temporal acuity, to the point of needing fairly long durations' (1987, p. 259). This dilated time provides a predictable experience for the listener and this allows them to focus on minimal sonic events (Imberty, 2005). However, this music is not frozen. It possesses an unconventional set of hierarchies (i.e., perceptual grammar), a sense of motion generated by forces in action and expectation, but it is not linearly goal-oriented: the ever-changing macro-structure incorporates dynamic and vectorial tensions (Imberty, 2005). Scelsi's *Quattro Pezzi* (1959) shows a sense of construction (for example, using the 'tension/resolution' schema) together with unconventional sonic meters (Menesson, 2008). There is a sort of expectation and structural hierarchy; the musical time is not linear, but a progressing sonic flow is discernible. There is a narrative but not a linear teleological profile: the listener is absorbed (or immersed) into a sonic stream. In other words, these music practices are founded on a series of essential sound configurations – based on frequency proximity, spectral similarity, comparable evolutions in acoustic properties – which mirrors a wide variety of experiential Gestalten of the physical world, and this – as we will see – is the key for an embodied engagement. Like all sounds, these configurations can engage echoic memory, STM, and LTM, but their particularity emerges in the working memory (WM).⁵ The primal grouping⁶ which often exceeds the time-limits of STM, the essential redundancy, repetition, and regularity which reinforces STM and opens WM up to 'semiactivated' portions of LTM (Snyder, 2001, p. 48), the lack of goal orientation and closures, the limited functional role of sound configurations and their formal independence from external causality, all these elements impact our memory – echoic, STM, and LTM – in different ways, essentially by strengthening STM and limiting the consolidation of LTM.

- *Echoic memory.* These pieces use sound types that emphasize their material nature. On the one hand,

they consist in massive pulsations, enveloping constructions of layered sounds, sonic clusters, all made of intricate textures. On the other hand, they explore the extremes of the audible spectrum, challenging listeners to deal with fragile sounds which lie at the threshold of audibility. Composers frequently exploit acoustic interferences based on the auditory phenomenon of critical bands and, as a consequence, due to the pervasive effect of aural beating, this sound gives the sensation that a concrete imprint has been left behind. Binaural beats, distortions, and sonic contrasts overtly show that to experience this music is to have a sense of a physical material that is stretched, cut, and expanded. The distorted sonic textures of *Triptych #2* (2008) by Verrando give the sensation that the music is being played (or emitted) very close to the listener's ear. When we listen to distorted blocks of noisy sound – as in Verrando's or Pan Sonic's music – we experience a sense of saturation, as if the sound material is physically filling a kind of 'container'. We have the sense of an acoustic pressure but also of restriction and a channeling of energy-forces. This kind of sonic impact is also revealed when sound disappears – that is, when the acoustic pressure of sound is relieved and the listener experiences a post-experiential absence marked in auditory sensory memory (Bonnet, 2012). A prototypical example may be found in *+/-* (1996) by Ryoji Ikeda. The Japanese sound artist uses sounds outside of the audible frequency range (e.g., in the piece entitled *-*, he adds to tones at 11 and 18 Hz) which leave a sort of unheard but felt sign.⁷ In particular, the last track of *+/-* consists of a very soft sound which is almost inaudible but which possesses an aura, a shadow of sound. Ikeda suggests 'a synthesis of opposites through its barely audible volume, short duration, and reliance on the faintest gossamer sound grains' (Demers, 2010, p. 87). In describing this last piece, he talks about 'a high frequency sound [. . .] that the listener becomes aware of only upon its disappearance' (Ikeda, 1996). Echoic memory of this perceived acoustic stimulus is therefore the key to constructing the identity of the sound.

- *Short-term memory.* In these pieces, primal grouping often goes beyond the time-frame of STM. A piece like *Pithoprakta* (1955) by Iannis Xenakis is conceived starting from stochastic principles and statistical laws, and its perceptual result challenges the listener to deal with contrasts of density and continuity rather than the succession of traditional tone-based constructions. Homogeneous musical sections often extend over one minute (e.g., bars 60–100; 123–170; 209–251) and are defined by their internal spectral coherence: a continuous sonic flow constructed on a micro-scale by entangled instrumental

voices (i.e., 46 solo string parts). Others contain long silences and detached sonic interventions that – due to the uniformity of the technique used (only *glissandi*, *pizzicati*, or *sul ponte*; bars 108–122; 187–207; 252–end) – define specific spectral ranges. In the piece *in vain* (2000) by Haas, the opening figure (being mirrored at the end) – which lasts over 5 min – recalls the never-ending development of some acoustic illusions such as the Shepard scale.⁸ The particular complexity of this episode favors the ‘bouncing mode’ of listening (Bregman, 1990): the intricate layering of glide patterns is heard following the principle of frequency proximity which chooses to group glides of the same spectral region. Thus, we are not able, during listening, to follow single instrumental lines but we engage a fluctuating sonic scene where distinct spectral regions are variably explored. Both Xenakis’s and Haas’s constructions interfere with any short-term grouping, but the continuous and repetitive profiles support listeners’ memorization of single patterns which continue to be restored in memory.

- *Long-term memory*. Similar to other music genres, schema-driven processes involving LTM draws on specific domains of experience. Several studies show how greater familiarity and competency favor an informed engagement: the typical public for CAM today is a sort of ‘new omnivore’ group made up of well-educated higher-status individuals whose likings and backgrounds embrace a broad set of musical practices (Grebosz-Haring & Weichbold, 2020). Trained listeners show a greater ability to discern instrumentation and deal with these uncommon sonic constructions, thus favoring mnemonic abilities (Wanke, 2019). However, the ‘memory sabotage’ of this music impacts physiological and more essential levels of perception that depends only to a small extent on training. The characteristics of this music restrict in fact formal-based *recollection* in LTM while exploiting the abstract nature of these sound configurations to favor LTM’s spontaneous processes (*reminding* and *recognition*).⁹ Listeners draw on familiar patters which concern essential forces and shapes identified in their general experience of the physical world rather than on consolidated musical recollections. If this music *sabotages* the construction of long stable memories, the schema-based segregation here emerges from the nature of the perceptual grammar itself: listeners grasp Gestalt and kinesthetic forces, in the sense that sound configurations are defined like one particular sensory dimension of some more general physical energy that could have been glimpsed with the other senses but which in this case is only present through sound (Bregman, 1990, pp. 401–402; Lehar, 2003). As we discuss later on, these forces pass internally to

our perception in the form of references to a quasi-physical realm (Fingelkurts et al., 2010). Schema-based processing recognizes sound configurations as profiles, shapes, or groups that unfold, expand, or collapse. The sense of this music concerns a series of forces and geometries – lines, planes, layers, degrees of density – that call in the physical world as potential: experiential Gestalten, kinesthetic patterns, textures, and immersive motifs that, due to their nature, lead to a non-sonic (but sonic-related) area made of intentions, processes, movements and appearances (Godøy, 1997; Wanke, 2019).¹⁰ This is to say that sound configurations such as contrasting sound masses, transient appearances of fragile overtones, or entangled descending lines are processed as energy-forms that we experience as impacting on us directly, and the schema-driven associations arising from these forces draw on our experiential and embodied cognition of the external world.

A piece such as *in vain*, for instance, by displaying a large variety of frequency-time configurations, is based on a continuous clash of harmonies and intonations (natural harmonic series vs. 12-tone equal temperament) (Hasegawa, 2015). Listeners may be unaware of the theories behind this acoustic relationship and still be captivated by the dynamic contrasts, the forms of the infinite descending movement, or the temporality of the progressive accelerations. But through these contrasting intonations, Haas also plays with our habits as trained or untrained listeners and our (un)certainties, such that one may wonder why horns and harp suddenly appear to be detuned or why some notes seem to come from another tuning system. These unnatural mismatches immediately offer the most engaged listeners the potential to construct through perception interactive relations with non-musical features by constructing high-level recollections with life events (which may recall contradictions or conflicts), or with corporeal and natural processes (in the form of frictions or the sense of loss), or even related to social conducts and cultural beliefs.

The characteristics (and our experience) of these pieces focus on the immanent aspect of sound, its characteristics and behavior. A potential criticism of this point of view is given by Erik Clarke, who claims that ‘[w]hen perception proceeds in an unproblematic way, we are usually unaware of the sensory aspect of the stimulus information and are only attuned to the events that are specified by stimulus structure. But when that relationship is problematic, the stimulus structure itself can become more evident’ (2005, p. 32). We argue here that it is precisely this kind of ‘problematic’ perception that arises when faced with today’s experimental music which allows some works in the genre to focus on immanent sensory aspects (the ‘stimulus structure’): the music of composers such as Haas, Lucier, Ullman, or performers such as Pan Sonic, Boards of Canada, or Thomas Brinkmann, is based precisely on unconventional sound

configurations that exploit sonic effects embedded within musical episodes, favoring a sensorial impact (cf. Kendall's phenomenal qualities (2014, p. 200)) and allowing the physical-acoustic energies of sonic stimuli to be central during listening.

We are not suggesting that an elaborated response does not exist or that this kind of music does not have in some way an emotional impact; rather it is that the best way to understand how we process this music is by thinking in terms of the mechanisms of perception we use in relation to sounds from the wider physical world. Our perceptual engagement with these pieces seems then to revolve around an embodied experience, a connection or bridge to the world in an everyday sense, more so than with other kinds of music where the knowledge of specific musical languages and cultural codes is central, and where the types of sound material cover a narrower range. It is not a complete novelty: studies (Bonnet, 2012; Clarke, 2005; Voegelin, 2010) examining the subject of perception in the context of CAM frequently draw on phenomenological perspectives and Gestalt theories (Koffka, 1935/1999; Merleau-Ponty, 1945/1976; Varela et al., 1991), elaborating on several disciplines within the so-called information sciences such as cognitive semantics, philosophy of art, and ecological approaches (Clarke, 2005; Gibson, 1979). The embodied experiences examined by these scholars reactivate certain concepts coming from early formalism but enrich the debate with a complex series of corporeal and environmental elements that expand the idea of music listening largely beyond its acoustic dimension. In going beyond Schaeffer's reduced listening (i.e., a listening attitude which focuses on an intrinsic sonic dimension and that disregards sound's production and origin), these authors elaborate theories that suggest the possibility of sounds holding a value that is either inherent to sound within a context (e.g., 'affordance' (Kozak, 2020; Windsor, 2000)) or that is part of wider poetic and esthetic spheres and which 'refers to the world beyond the musical work' (Demers, 2010, p. 36) (e.g., 'Intention/Reception' project (Weale, 2005)).

Given the nature of its properties, this music is particularly apt to undergo an early stage perceptual segregation based on the same principles (Gestalt and kinesthetic) of which it is made. Even if it is true that perceptual segregation is central to how we perceive all music (given it is fundamental to perception), it seems that this music suits this type of perceptual processing particularly well. This music appears to be specifically grounded in the basic essentials of auditory perception rather than in articulated culturally-based musical forms and figures. These elements are actually time-spans, or 'temporal Gestalt-units' – as they are defined by James Tenney – 'whose perceptual boundaries are largely determined by the nature of the sounds and sound-configurations' (Tenney & Polansky 1980, p. 205; see also Reybrouck, 1997, p. 65). A piece such *in vain* by Haas, for instance, displays several forms of

repetition (Massoud, 2017) exploring a limited set of sonic features (Varga, 2011, p. 106) and these aspects certainly constitute 'a kind of implicit rehearsal that eases memory load' (Snyder, 2001, p. 52). Nonetheless, it includes musical episodes (which correspond with the extinguishing of the lights in the performance space and that are unmetred in the score) that are characterized by a sequence of sustained sonic aggregates (using both tempered and natural tunings) which merge into a unique sonic flow with indistinct temporal indications. These episodes – as mentioned above – extend beyond STM's time-frame, but in lacking teleology and functional organization, hamper the construction of long-term memories of sequential information. More generally, characteristics such as redundancy, repetition, and source ambiguity challenge our perceptual approach. The early stage of perception seems the crucial level to examine in order to understand how we listen to this music and – more specifically – how we organize and recall these sounds in a meaningful way. It appears evident that the link between memory and this kind of music needs a broader perspective capable of encompassing neurological and psychoacoustic models of our early stage of perception and opening up to different perceptual stances and expectations.

Perceptual Pathways

Academic enquiry into music and sound perception tends to follow one of two main approaches: on the one hand, analytical assessments mainly in psychology and neuroscience tend to explore sonic qualities through empirical testing and to promote consistency in the research (Giordano & McAdams, 2010; Grey, 1977). On the other hand, philosophical and semiological scholarship conventionally refer to the Western classical repertoire, world music, and popular music (Davies, 1994; Kivy, 2002; Ridley, 2004). Within this large group of studies on music perception, only a small number deal with CAM moving across these two different approaches. Some studies tackle this research using theoretical premises (Terrien, 2012; Windsor, 1995), while others use bottom-up empirical methods, employing surveys or subjective resources (Frey et al., 2009; Glover et al., 2018; Weale, 2005). A number of studies, which deal with perception of atonal music (Dibben, 1999; Mencke et al., 2019; Ockelford & Sergeant, 2012), inaugurate a research-based potential toward 20th- and 21st-century music that is very promising. However, within the large panorama of experimental music practices, classical atonal music occupies a limited portion. This, in particular, in continuing to be largely based on 12-tone equal temperament (12-TET) invents new note-based approaches beyond tonality. The key aspects of the music considered in this paper come from those music practices which cut across 12-TET rather than just going beyond tonality (e.g., inharmonic spectra, microtonality, noise, concrete sounds, psychoacoustic phenomena). The lack of perceptual studies

on this music seems paradoxical as many genres in this context are themselves concerned with these very questions of sonic perception, but the reasons for this are to be found in the peculiarity of this music in challenging the listener with a wide range of sounds and evolutions against the more restricted set of conventional musical sounds.

The music we consider here does not use established cultural codes and, as yet, there is no established repertoire, but it is still presented to the listener as music. At the same time, the material used includes many sound shapes and forms that appear to reflect everyday sounds. So, the challenge for the listener and the analyst is how to understand and describe this music that does not use established musical forms, but equally is not simply a catalogue of field recordings. Does perception operate in a particular way here? We propose it does. In examining in detail our auditory process, our contention arises from the belief that the range of sounds used in this music changes the way we should look at and understand the role of these early stages of perception and their importance. We map out an overview of these stages in order to tie in neurophysiological processes and memory steps.

Stephen McAdams, Tim Griffiths, and their colleagues (Egermann & McAdams, 2013; Griffiths & Warren, 2004) isolate a first step of encoding or transduction in which our perceptual system reacts within a psychophysical dimension and carries out a primary grouping of sounds identified by Albert Bregman as primitive segregation (1990) and by Michael Kubovy and Van Valkenburg as the early perceptual stage (2001). Within this early stage, lasting approximately between 250 and 500 ms (Koelsch & Siebel, 2005), a series of decoding steps deal with sound features starting from the most essential ones, such as event fusion (echoic memory) and pitch and temporal cues (continuous vs. transient), to arrive at more complex sorting tasks that include harmonic and rhythmic grouping (STM) (Koelsch & Siebel, 2005; Norman-Haignere et al., 2015; Snyder, 2001). What seems to be clear is that a considerable part of the segregation and matching of spectro-temporal patterns, i.e. the auditory scene analysis and auditory stream segregation (in particular the formation of auditory Gestalten which entails the processes of melodic, rhythmic, timbral, and spatial grouping), occurs in the primary auditory cortex (Koelsch & Siebel, 2005) and largely involves the STM. Once the sonic message is segregated, the next steps correspond to Bregman's *schema-based segregation* which activates the knowledge of familiar patterns stored in LTM. At this stage, the perceptual attention leads to a figure/ground segregation (Kubovy & Van Valkenburg, 2001), thus allowing the activation of abstract knowledge structures (McAdams, 1996). From this stage, psychoacousticians propose the activation of induced emotive mechanisms, those related to episodic memory (implicit and explicit) and the initial representation of musical forms and complex expectancies (Juslin et al., 2010).

The Vertical Stance

The juncture in perception where memory starts to take on a fundamental role coincides with the formation of Gestalt structures. We advanced the idea that our primitive segregation of an acoustic stimulus is, in some ways, connected with our ability to store (memorize) and predict stimuli, and that our image schemata emerge from a set of frequency vs. time constructs (Fingelkurts et al., 2010; Griffiths & Warren, 2004). When we consider the sonic characteristics of specific styles within CAM, we believe that the integrated process of segregation, storage, and prediction of upcoming stimuli becomes a powerful function of our auditory system. This process is performed by the WM, which allows the performance of STM and semi-activated portions of LTM to be connected together in the present (Snyder, 2001).¹¹ A recent study suggest that our WM deals with unified *auditory objects* rather than separated acoustic features (Joseph et al., 2015), and the sound configurations, described above, appear to correspond to these fundamental perceptual units as coherent configurations in which acoustic components are segregated by spectro-temporal cues such as common onset, frequency range, and modulation. Moreover, the lack of articulation and limited dialectic in this music, the global unicity and variety of sound types, the systematic repetition, the non-teleological construction, all these factors may entail a particular functioning of WM.

While STM is strengthened (rehearsal, continuity, coherence), the semi-activated LTM is involved as a platform to construct progressively higher-level associations. More importantly, the combined process of segregation/storage/prediction leads to the formation of a kind of stable matrix – generated by the periodicity of repeated patterns or continuous textures – combined with fluctuating elements (i.e., sonic transients, combination tones, minor interferences and resonances, noises, distortions, all the way to minor psychoacoustic phenomena). By incorporating unstable sonic elements, this stable matrix, a vertical 'now',¹² allows for both a full predictability and a complete openness to unpredictable events.

We associate this specific perceptual stance with a vertical temporality in perception (Kozak, 2020; Kramer, 1988). While Kramer's idea of verticality in perception relates to a lack of hierarchies and an absence of musical discourse, Kozak recently advances the idea of a vertical stance that belongs to a personal perceptual state and is seen as a flow of consciousness which moves vertically (orthogonally with respect to linear time) (2020, p. 238–247). This perceptual stance with respect to music clearly applies to spectralist and minimalist perspectives. Bergson also gave an account of this idea of an inner time that is non-linear. His notion of duration applies to a music that is an expression of time and where the musical material is a manifestation of pure duration. This description brings to mind the work of Grisey.

The innermost temporality of Grisey's music, inspired by Bergson's intuitions, is realized through a distortion of a

sound's perspective in memory (Hervé, 2004). Memory, for Grisey, is the locus of the *mental resonance* of sound: a sort of re-exposition of the trace of sound: 'through to its retroactivity', memory becomes the way in which we 'rebuild the trace described in time by what lasts' (Imberty, 2005, p. 53). Music manifests itself as pure duration, or – seen through Grisey's eyes – '[music] gives us to be perceived what Proust called "a little time in a pure state"' (Grisey, 1987, p. 269). In this perspective, WM is the locus of the mental resonance of sound, a re-exposition of the sound, within this everchanging matrix, allowing us to perceive a sound's identity and continuity. The chronometric time (i.e., linear time) of these pieces is then transformed through the process of segregation, storage, and prediction into a sort of personal dimension of consciousness in which this *resonant time* provides the domain in which sound's essential features exist, entering or exiting our experiential frame. This vertical stance, which establishes itself through a shift from a real-time perception of music to a post-experiential condition, is characterized by the constant link between external and internal temporalities.¹³ During listening, the transition between external (chronometric) and internal temporalities involves in fact a long chain of events which begins with elemental physical operations that have a non-phenomenal and purely neurophysiological nature. These operations are 'local fields' consisting of neural activity synchronizations (i.e., electrical current patterns of amplitude, phase, spatial and temporal rates of change) (Fingelkurts et al., 2010). Successive synchronizations of these 'local fields' lead to complex brain operations. These – by holding both neurophysiological and phenomenal/subjective ontologies – could be mutually synchronized within a *new* time-scale, leading to abstract and more complex spatio-temporal patterns which are at the base of integrated phenomenal experiences (ibid., 2010).

Due to the specific traits of this music, listeners are in fact able to construct a *continuous instant* in perception – based on the matrix defined above – where experiential Gestalten and kinesthetic forces occur. This continuous instant is periodically fed by new recollections and anticipations: memory glues together retentions, real-time experiences and forthcoming events into a suspended ever-changing temporal domain. This spectro-temporal domain of perception is accessed within this vertical stance.

The Internal-External Framework: A Morphodynamic Analogy

In listening to this music, we perceive an external acoustic stimulus that goes beyond its source and spatial location and reifies itself within our internal perceptual domain. When listening to Haas' *String Quartet n°2* or to Lucier's *Two Circles*, these sonic sensations seem to concern something more complex than can be accounted for simply in terms of source production (Wanke, 2016). The sound originally produced by the instruments becomes a sort of

perceptual entity independent from any apprehension of source. The actual source seems indeed as if it is not sufficient to account for our experience of the sound, a sound that exists in the domain of our perception. This is due in part to how the sounds used seem to lack external spatial cues, how the sounds are formed over time, as well as the continuous shifting between the reality of external sonic experience and the nature of the perceived sound configurations of the perceptual grammar. This grammar mirrors the forces and shapes of the external world but crucially does not include elements relating to external space (sound origin, directionality, source position).

This is the key characteristic of the sound configurations present in this music: sounds in these pieces, coming from either acoustic instruments, electronic or natural sources, at first bear both intrinsic features and external information relating to the actual origin of sound. When we approach this music 'vertically', the portion of sound concerning external cues (interaural level and phase differences) is naturally processed and then left behind, allowing us to better focus on the internal sonic cues. The elements of sound belonging to the perceptual grammar pertain to the intrinsic spectro-temporal properties of sound which disregard sound's external origin and position. When we listen to a descending *glissando* in Haas' *String Quartet n°2* (1998), to an opposition of sound masses in *Ilmenemismuoto* (2004) by Pan Sonic, to fluctuating interferences in Lucier's music, or to the fragile evolutions of Ullman's pieces, we grasp these sonic figures separate from any search for an actual cause or concern for the movement of a sonic source: we access a sonic domain in perception.

Listeners grasp this domain made of profiles, shapes, and groups that unfold, expand, and collapse, and articulate their perception of sound in terms of lines, planes, layers, and degrees of density (Godøy, 1997; Wanke, 2019). This perceptual domain, we argue, is a sort of conceptual extent where the sonic geometries of the outer world-sounds are internally reproduced. Within an internal-external framework, the sound configurations are transduced internally, from physical to electrical energy, through the processing of our perception in which the 'cochlea constructs sound images with dimensions that correspond to time and frequency domains' (Griffiths & Warren, 2004, p. 888). The Gestalt forms described here correspond to the internal Gestalt process of segregation, storage and prediction. During listening, we process the external acoustic features of sound – which can be effectively represented on a spectrogram – and we take in these features internally as a sort of 'perceptual spectrogram'. This parallel, a sort of morphodynamic analogy, between an *external* spectrogram and an *internal perceptual spectrogram* is precisely based on this spectro-temporal processing that our memory stores and predicts, a region defined by spectro-temporal dimensions and not spatial ones.¹⁴ The perceptual spectrogram is a domain, as defined by Peter Gärdenfors (2000), 'a set of integral

dimensions' where these (quality) dimensions assume a key role in representing concepts.¹⁵ A 'high pitch' is not literally high in our world, but neurophysiology tells us that a physical distance between high and low frequencies exists in hair cells along the tonotopic axis (Shamma, 2001). Therefore, to what extent does the mental representation of the descending *glissando* in Haas' *String Quartet n°2* as a descending profile depend on a metaphorical construction or a physical distance between stimuli? While the vertical representation of pitch distance seems to not be congenital (Eitan et al., 2012) and while in other musical cultures this frequency dimension is described as an axis of 'sharpness/heaviness' or 'small/large' (Eitan & Granot, 2006; Zbikowski, 2008, 2017), there is a common perspective in conceiving the separation between frequencies as a kind of extent in our perception and certain studies seem to advance suggestions in this sense (Fingelkurts et al., 2010; Gallistel, 1990; von der Malsburg, 1999). Several authors discuss the metaphorical descriptions of music listening, mainly accounting for the potential of these figures of speech. Snyder, for instance, reports the sensation of motion provided by a sequence of notes (i.e., stable pitches) as an *illusion*, but sound configurations such as *glissandi*, binaural beats, and microtonal cross-fading transitions concern real movements in the dimensions of frequency or intensity. Therefore 'if phenomenal consciousness is a biological phenomenon within the confines of the brain, then there must be some specific level of organization and some specific spatial-temporal grain in the brain where consciousness resides' (Fingelkurts et al., 2010, p. 217).

In trying to match the brain and mind's spatio-temporal dimensions, Fingelkurts and co-workers discuss in fact the presence of an operational space-time in the brain that is *functionally* isomorphic to the phenomenal space-time in the mind (2010). The isomorphism is not of a first order (i.e., an internal neuronal process that corresponds to the experience of an external stimulus of a 'circle', for instance, does not involve neurons arranged in the form of a circle) but rather is functional. The crucial shift here is between 'sharing physical properties' and 'mirroring functional relationships', that is to say that neurological and experiential systems consist in the different realization of the same kind of process. While Lakoff and Johnson talk about 'orientational metaphors' in the case of language (2008, p. 14) and Godøy puts forward the *shape-paradigm* in the discussion of music cognition (1997),¹⁶ we find appropriate for the definition of these internal geometries the notion of 'image schemata', which by having an inherent spatial structure and a kinesthetic character, are able to bridge our perception to the external world within an embodied experience (Johnson, 1987).

Our hypothesis here is that this music appears particularly capable of prompting the formation of these 'internal geometries' which are essentially temporal-based rather than spatial-based (i.e., they incorporate dimensional extent

but do not hold specific external spatial information). In line with Kubovy & Schutz's thesis (2010), our contention is that the intrinsic properties of sound relate to temporal, spectral and dynamic dimensions, while spatial elements do not pertain to this constitutive portion of sound. This perceptual domain – accessed through the vertical stance and realized by WM – is characterized by the following aspects:

- it exists within a precise experiential frame (typically during listening and acoustic recollection);
- it can leave this experiential frame and return in memory, thus we grasp its presence and absence;
- it originates in a low-level physiological spatial delimitation (along the tonotopic axis of hair cells in the inner ear) which is preserved in terms of topographic distributions in the brain (according to high- and low- frequency spectral occupancy, transient and continuous characteristics, (Norman-Haignere et al., 2015)) and is functionally reproduced in this perceptual domain;
- this topographic map in the auditory cortex defines a domain – in the sense of cognitive semantics – which is constituted in terms of integral dimensions. This domain may present a kind of inner dimensional functioning (vectorial)¹⁷ and is aspatial, that is it does not hold specific cues from external space.

Given this, a perceived *glissando*, i.e. an effective continuous motion in the realm of frequency/time/intensity, may originate a set of responses in our brain that are reproduced in this perceptual domain as a functionally coherent set of conscious references. The functional isomorphism between brain and mind (Fingelkurts et al., 2010) allows us to hypothesize that these references elicited by the activity of listening to a *glissando* preserve essential elements of its nature such as those of wholeness, continuity, directionality, movement, and progressive change. This domain is then embodied domain in our memory, a pure temporal realm of auditory consciousness that is aspatial. In order to give sense to the apparent paradox in conceiving of a physical domain in our brain, defined by time/frequency/intensity dimensions but that holds an essential aspatial nature, we draw on the analogous concept of the No-Space world formulated by the philosopher Perter Strawson.

The World of Sounds and Memory in Strawson's No-Space World

We have seen how certain musical practices of CAM are based on a set of Gestalt features which are processed by our auditory system in a sequence of operations within a realm that exists as a reconstructed spectro-temporal domain of our perception. But what do we preserve as the essence of a sonic experience in memory when we remember a sound stream?

We can reformulate the question as follows: where do sounds go when they exit the frame of experience? The way in which we experience an entity ceasing to exist is different from the way in which we experience our inability to keep perceiving it. There seems to be such a distinction in our ordinary experience (consider the difference between seeing something blurry and seeing something evaporating or a stone falling into water vs. a drop of water falling into the sea). Can we assume that a similar distinction also holds in a world of ‘pure events’ – like the musical one – or in what Peter Strawson called a *No-Space world*?

Strawson conceives of the No-Space world as a merely auditory world, possessing only a temporal dimension and no spatial dimensions.¹⁸ He treats sounds as a paradigmatic case of entities that may appear to exist solely in the temporal dimension. He wonders if there really can be merely temporal entities or if, instead, even when an entity seems to be merely temporal, there is some underlying spatial dimension that has to be made explicit. Strawson’s No-Space world thought experiment rests upon the assumption that we can conceive of an experience of sounds lacking a spatial dimension, and this argument appears to match with the idea of an aspatial domain in memory achieved during a perceptual vertical stance. Although most sounds are experienced as involving a certain distance or direction, there are cases, such as listening to music through headphones or – as mentioned above – experiencing acoustic illusions (binaural beats) (Wenzel, 1992), in which the auditory experience seems to be purely temporal and to not have any spatial component. In particular cases, when listening to music, the more we set aside information relating to the external world whether acoustic (directionality, localization and spatialization) or visual (when closing our eyes, our listening capacity seems to be sharpened), the more we feel able to grasp the intrinsic matter of sound (Degerman et al., 2006; Farthofer, 2007; Ross, 2010).

For Strawson’s hypothesis to hold, it is sufficient to admit that we do not individuate sounds (only) spatially, or that they are only contingently spatial:

Sounds can maintain their identity even when they lose their spatial properties [...] You can identify and selectively attend to a particular sound non-spatially. [...] We can hear a sound without hearing it to have any spatial property at all [...] We can imagine the actual world of sounds as a no-space world (Nudds, 2001, p. 215).

According to Strawson, the basic notion of objective reality, which we will hereafter call thinghood, involves the existence of self-standing re-identifiable individuals. These individuals exist independently of any experience of them. Yet, some of them are, and others – in suitable conditions – would be objects of experience for a subject, who in principle can re-identify them in distinct experiences occurring at different times. Re-identification on the part of the subject rests upon the fact that an individual

remains the same not in the sense of mere qualitative sameness, but in the much stronger sense of numerical identity. As Swanson puts it in his discussion of Strawson’s thought experiment, ‘this is the crux of the notion of thinghood’ (1967, pp. 223–224).¹⁹ The main question that Strawson addresses in Chapter II of *Individuals* is whether this notion of thinghood can be satisfied by entities existing only in time. He considers sounds whose experience involves a temporal succession but not a spatial ordering. As merely temporal individuals, sounds are characterized just by their frequency and loudness, as well as by their position in the temporal succession.²⁰ Such an account of sounds leads Strawson to conceive of a merely auditory world that could do without spatial ordering, namely a ‘No-Space world’ (1959, p. 63). This world is inhabited by audible sounds, and by at least one subject of experience who can hear sounds that exist only in time – a pure hearer whom, following Evans (1980, p. 276), we will call ‘Hero’.

If the No-Space world has thinghood, then Hero should have the possibility of ‘identifying a particular sound as the same again after an interval during which it is not heard’ (Strawson, 1959, p. 70). Without the idea of existence as being perceived at times and unperceived at other times, then indeed, perceived objects cannot be detached from the experience of them, and therefore there cannot be any idea of thinghood as an objective reality.

In our spatiotemporal world, re-identification is possible because the experience of space involves a distinction between what is inside and what is outside with respect to the spatial frame of experience. By providing locations in which things can also exist beyond the frame of experience, space allows things to exist unperceived, thereby having what we will call, in analogy with an off-stage situation in theater, an *off-experience existence*. Still, if the only available dimension is time, as seems to be the case in the No-Space world, then both Hero’s experience and all sounds will occur in that unique dimension, and therefore there seems to be no room for an off-experience existence. If a sound is in time when Hero’s experience occurs, Hero necessarily hears it; if it is not in time, it does not exist.

In order to have thinghood also in the No-Space world, we would need a dimension capable of containing potentially perceivable but actually unperceived individuals. A sound in the No-Space world must be able to exit from the frame of Hero’s experience, to exist for some time while unperceived, and then to possibly re-enter into the frame of Hero’s experience. For this purpose, we would need ‘the necessary non-temporal dimension for, so to speak, the housing of the objects which are held to exist continuously, though unobserved [...] It seems we must have a dimension other than the temporal in which to house the at present unheard sensory particulars, if we are to give a satisfactory sense to the idea of their existing now unperceived, and hence to the idea of re-identification of particulars’ (Strawson, 1959, p. 74).

Strawson addresses this problem by introducing the notion of master-sound, i.e. a sound that is always audible

by Hero and that has its own peculiar timbre, constant loudness, and varying pitch. In the No-Space world, the master-sound determines the frame of Hero's experience. At a certain point in time t , Hero can only experience the sounds that, at t , are in the proper relation with the master-sound's pitch at t .

It seems plausible that, if Hero simultaneously hears two distinct melodies, he can identify them as distinct individuals simply by virtue of their intrinsic features of pitch, loudness and timbre. Therefore, the non-temporal dimensions needed in order to have thinghood are frequency and intensity. In this sense, Hero can simultaneously experience distinct individuals even if he lacks the direct perception of a spatial distance: he does not need a spatial relation to the sound/individual but only a time/frequency/intensity world.

This claim can be strengthened by considering a series of empirical cases. Kubovy (1981) outlined the so-called 'Theory of Indispensable Attributes' (TIA), which hypothesized the indispensable attributes that are necessary for a perceptual object to be such. In the years when psychologists and neuroscientists – followed by philosophers – began to pay attention to modalities of perception other than vision, Kubovy was the first to actually speak of auditory objecthood, by extending the TIA to audio and audio-visual objects. As Kubovy and Van Valkenburg put it: 'Notions of objecthood have traditionally been cast in visuo-centric terminology. [...] we re-examine the concept of an object in a way that overcomes the limitations of the traditional perspective. We propose a new, cross-modal conception of objecthood which focuses on the similarities between modalities instead of the differences' (2001, p. 97). By presenting to an observer two spots of light on a surface, for example, Kubovy & Schutz (2010) first demonstrate that spatial separation is an indispensable attribute for vision. Both spots of light are yellow and they coincide; in this case, the observer will report one light. Now suppose the color of the lights is changed by the experimenter, so that one spot is blue and the other is yellow, but they still coincide; in this case, the observer will report one white light. For the observer to see more than one light, the spots of light must occupy different spatial locations. In an analogous manner, pitch separation is an indispensable attribute for sound. Imagine playing two simultaneous 440 Hz sounds for a listener. Both of them are played through the same loudspeaker; the listener will report hearing one sound. Now suppose we play these two sounds through two loudspeakers; the listener will still report hearing one sound. For the listener to report hearing more than one sound, they must be separated in pitch. We can thus draw the following conclusion: in audition pitch is an indispensable attribute for perceptual numerosity whereas, unlike the case of vision, space is not. After all, the manifest independent processing of spatial and non-spatial properties of acoustic stimuli in WM (Alain et al., 2001; Furukawa et al., 2013) is evidence that the identification of

sound (pitch, spectral and temporal issues) and spatial location (based on interaural time differences) are – at a primal stage – retained and processed in separate streams. Moreover, Alain and his colleagues demonstrate to which extent processing sound identity and sound location depends on diverse, specialized auditory pathways within the same individuals. In particular, processing pitch information recruited brain areas distributed in the ventral part of the brain, whereas sound localization recruited areas distributed in dorsal regions. The conclusion is that 'the neural systems involved in identifying and localizing auditory objects are functionally and neuroanatomically segregated based on task demands even when stimuli are identical across tasks' (2001, p. 12305). Their results suggest that the 'ventral and dorsal prefrontal areas are involved in representing two distinct types of information about the environment, "what" and "where"'.

Applying these insights to the relation between a real-time experience and the post-experiential condition (memory) of a musical episode, we can see how a particular listening experience – that we call the vertical stance – of certain currents of CAM offers to the listener the opportunity to grasp the gist of sounds. By processing *auditory* objects, the WM is the nexus of convergences of these very different perspectives. The crucial step here is the introduction of the term 'auditory event' (Blauert, 1996) to describe the subjective perception of a subject listening to a certain situation. The definition was also meant to clearly distinguish between the physical sound field and the auditory perception of the sound: 'Terms such as "sounds source", "sound signal" and "sound wave" will always be used to describe physical phenomena that are characteristic of sound events. What is perceived auditorily will be denoted by the adjective auditory, as in the term "auditory object", or, preferably "auditory event"' (Blauert, 1996, p. 3). This distinction matches the external-internal framework described above where the morphodynamic analogy between external acoustic features and internal auditory objects is realized.

Seen from this perspective, the sound configurations outlined earlier represent the very essence of sound, i.e. refined from data coming from its effect in the physical space. These sound configurations are essential components retained in memory which may (or not) enter our experiential frame. Conscious of the fragility of this cross-disciplinary parallel, we put forward the idea that our recollection of sound in memory can be nothing but the re-identification of parts of several process-like individuals. Given that the spatial features of an incoming acoustic stimulus are a *non-intrinsic* property of sound, the recollection of sound in memory is essentially the reification of *pure* sound itself without extrinsic information. If our listening is an articulated (often multimodal) process in which frequency, dynamic, temporal, and spatial features are processed, then the 'offline' recalling process of memory

occurs in a time/frequency/intensity domain, the Strawson's No-Space world.

Being characterized by a set of integral dimensions which guarantee the thinghood, Strawson's No-Space world, the memory, is the domain where Hero may experience auditory objects without external spatial references. Memory, seen as containing non-spatial representations of sounds, is in fact both experiential – in the sense that is capable of exciting emotional, motor-programmatic, and even linguistic dimensions – and a housing for sounds outside of experience (Grimshaw & Garner, 2015; Zatorre & Halpern, 2005).

From an Aspatial Domain Towards a Direct Embodied Experience.

Working memory is a crucial meeting point between the philosophy of perception and several studies in cognitive science. We embarked on this discussion by starting from specific musical characteristics and empirical observations (Santarcangelo & Wanke, 2020). The unique features of certain music practices today challenge analysts and scholars and encourage researchers to propose new perspectives to approach perceptually a kind of music that is still largely considered an isolated (and even elitist) form of musical expression. We have seen how some authors of CAM develop a musical discourse based on intrinsic sonic characteristics, thus making use of physical acoustic phenomena to compose their pieces. Moreover, the impact of this music is often radically embodied, in the sense that it relates to corporeal and material facets of our experience, such as in the immersive sonic experiences of sound art and electroacoustic performances, or with concerts of noise music and electronica where the audience is faced with the physical materiality of sound. In order to account for this embodied engagement, we present the idea that the potential of this perceptual grammar is realized through a vertical perceptual stance within an internal-external framework.

In a first step, from external world to internal perception, Gestalt configurations of the perceptual grammar are processed as image schemata in a domain which reflects the external dimensions of sound (time, frequency and intensity). While this grammar does not include spatial information relating to sound types and origins, the image schemata in such domains depict information that is somehow dimensional (presence/absence, stasis/motion, ascending/descending). The difference is easily grasped when we compare the types of kinesthetic information provided by a descending *glissando* or a sound of a car passing by. Essentially, the *glissando* gives rise to a sense of spatial movement that occurs internally: our relation to it is not spatial, as it is in the case of the sound of the car. As in the 'No-Space world', we do not relate to a *glissando* in terms of a certain positioning away from us, rather these sound configurations express tensions, forces, and movements that seem to occur directly to us.

Indeed, a second step, from internal perception toward the external world, should also be considered as it realizes the idea of embodied experience. The perceptual grammar of this music is determined by Gestalt categories with hierarchies and interrelations, and is made up of basic and structured principles, motions, and general forces. This field of forces mirrors tensions and movements of the external physical world. This apparent conflict – the lack of external information of sound vs. the connection to the experiential world through the presence of essential forces in sound – is the essential characteristic of this music: a music that explores sonic forces which refer to a quasi-physical realm. The listener apprehends not just a sonic material but also a potential physical energy-form. The experience of sound then pushes beyond the sonic to a potential concrete form. To listen to this music is to search within a personal sense of physicality and relation of perceiving body to the world, a relation made of tensions, motions, contrasts, and recognitions.

This is the crucial step from the aspatial perceptual domain towards the embodied experience. These sound configurations are in fact of a dual nature: abstract and concrete. On the one hand, by tying in with the auditory Gestalten processing at the early stages of perception, they detach themselves from their origin, thus they are abstract and do not hold figurative elements of the world. On the other hand, precisely due to their intrinsic sonic nature, they pass into perception as energy-forms: they discard the superficial appearance of things and go beyond a straightforward cognitive abstraction or any elaborated constructions, being able to instead reproduce these forces directly. The potential prompted by these sound configurations represents both a vivid concreteness and a lack of precise identification, thus leading to a set of associations which are largely unpredictable.

In other words, the listener of this music, facing a sonic substance in motion, is forced to reflect and make sense of these aural forms which do not point to anything specific in the world (i.e., aspatiality) but convey general tensions, forces, and impulses of the world. This situation encourages not just a reflection on the experience in the moment of a sound but draws in a more complete consideration involving that particular listener's embodied identity in the world, their experience in memory and context. Therefore, a descending line moving in this aspatial world of sounds acts immediately on our body giving us, for instance, a sense of falling. This is due to the fact that this sound configuration reproduces an energy-form as happening to us, and the sense of falling is nothing more than our experiential response to this impulse. In this sense, our primal associations are embodied as they meet a demand for a link to our bodily experience. Sound configurations here are *auditory events* in the sense that they are prone to a subjective embodied engagement. These configurations tie in with the wider area of our perception of the world as a mass of material energies. Within the No-space world,

there is not a self-projection of the Hero in a kind of spatial relation to the sound, but rather this domain is directly embodied thus allowing sound configurations to impact directly on the Hero.

We can see therefore this grammar as made of ‘bridging elements’ for their *potential* in reproducing the most fundamental set of external forces in our perception. The potential internal-external link is what allows embodied experience to handle these Gestalt configurations in this way, even if it is not linking them to a specific external origin. The bridging elements of this music open up to a listener’s experience of the world and constitute the means to access the non-sonic dimension made of high-level cognitive associations (that we have only mentioned here).

Conclusions

The rigid pulsations of *Etäisyys* (2004) by Pan Sonic or the intricate descending line traced by Haas in *in vain* are sound configurations: they hit our ears and capture us through the acoustic pressure. They leave geometries, shapes, motions, and tensions, and all these acoustic traces convey a presence and a set of gestures that we assimilate and integrate within and according to our diverse backgrounds, enriching these with our experience in a continual relationship between the sonic and the personal. What differentiates this interaction between sonic perception and a broader cognitive apprehension from the processes at work in the perception of more traditional music is that the kind of music discussed in this paper deliberately seeks out and opens up this territory while still remaining grounded in a certain kind of sound. This is made possible through the unique characteristics of this music, its perceptual grammar, the particular processing of its Gestalt concatenated sonic structures, the way it is presented, and how it pushes listeners to assume a particular perceptual stance.

Some time ago, just after a concert of British electronic duo Emptyset, a friend of us asked: ‘Why, when I listen to this music, do I think of scrap metal and wish to break everything?’ The embodied experience of this music appears to be fully realized through its sound configurations. These seem to detach themselves from the actual origin of the sound in external space and are processed and retained in memory while still having an implicit connection to the world, namely in the nature of constituting forces and energies. By lacking spatial information of the external world, these forces and energies hold the potential to access a large variety of high-level cognitive responses impacting directly on the perceptual self, the Hero. This kind of potential, whether contained in a suspended frequency interaction, a descending spiral profile, or a dense alternation of sonic aggregates, leads to an unpredictable set of high-level associations. This is why when we listen to continuous descending *glissandi* or a massive contrast between distorted and almost imperceptible sonic situations, we realize this extra-sonic potential in sound by

engaging in high-level associations, experiencing for instance a sense of falling down a cliff or the mental imagery of solid blocks.

We have shown the multifaceted connection that may be found between certain genres of contemporary art music and memory. It is not a question of mere mechanisms of memory recollection, but rather that listening attitude and memory take a fundamental role in the perception of this music. This function is so intense that we may equate auditory memory with a time/frequency/intensity domain in which this music reifies itself in our perception. The experiential Gestalten of this music may be efficaciously represented as a three-dimensional domain of our inner perception of sound. This domain is without extension, that is, without information of the external space (sound origin and directionality), but with a *potential* to construct a bridge toward an embodied engagement. From this point of view, our perspective may open up the possibility that the ‘No-space world’, the physically impossible idea in the world, actually exists in our memory.

Peer review

Lawrence Zbikowski, University of Chicago, Music Department. One anonymous reviewer.

Lindsey Reymore, McGill University, Schulich School of Music.

Action editors

Amy Belfi, Missouri University of Science and Technology, Department of Psychological Science.

Kelly Jakubowski, Durham University, Department of Music.

Contributorship

All authors researched literature and conceived the study. All authors wrote, reviewed, and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Authors received funding for this research by CESEM - Nova University of Lisbon and Fundação para a Ciência e Tecnologia (FCT), Portugal (grant number UID/EAT/00693/2019) (R.W.); and by University of Bergamo, Italy (grant: STaRs Supporting Talented Researchers 2019) (V.S.).

ORCID iD

Riccardo Wanke  <https://orcid.org/0000-0002-5480-7905>

Notes

1. We use the term ‘contemporary art music’ to refer to a large set of current genres and esthetics, both institutional and underground, that are generally included within current

- experimental music practices. This is an impartial term that takes the place of other common terms such as ‘new music’, ‘contemporary and experimental music’, or ‘avant-garde music’ (Grebosz-Haring & Weichbold, 2020).
2. We refer here to the famous definition of Edgar Varèse and Wen-chung of music as ‘organised sounds’ (1966).
 3. ‘des couleurs, des effets de lumière, des sensations tactiles, des propriétés des matières physique’ (Bériachvili, 2010, p. 12, our translation).
 4. E.g. Michel Chion’s relation between inner and outer musical space or Annette Vande Gorne’s distinction between real, designed, and virtual spaces. For an exhaustive description please refer to Chouvel & Solomos (1998) and Solomos (2019). In the philosophical literature, the relation between auditory perception, its objects, and spatiality is highly debated (see, for example, Nudds, 2001; O’Callaghan, 2010; and, for a review, O’Callaghan, 2020).
 5. The working memory is a definite portion of memory which includes STM and portions of LTM (Snyder, 2001, p. 48) and is the crucial system for the maintenance and manipulation of information, taking on an active role and being characterized by a mutual functioning of storage and processing. Recent findings put forward the idea that WM is the emergent property of a multiregional network in the brain (Fingelkurts et al., 2003).
 6. This is the very first process of segregation which takes place at the ‘bottom’ of the perceptual-cognitive system and uses basic and largely involuntary processes that are therefore referred to as ‘bottom-up’ and have been formed through the evolution of our sensory systems (Bregman, 1990).
 7. These are inaudible infrasounds which may provoke neural entrainments, organ resonance effects, nausea, physical impact, and respiration inhibition (Goodman, 2010, p. 11).
 8. The initial descending pitches are tritones, G#6/D6/G5/C#5/G4/C#4/F#3/C3, while when the section reappears (bars 344–413), the descending line consists of an almost chromatic series.
 9. ‘[R]ecollection [occurs] where we intentionally try to cue a memory; reminding, where an event in the environment automatically cues an associated memory of something else; and recognition, where an event in the environment automatically acts as its own cue. Recognition and reminding are spontaneous processes that are going on constantly’ (Snyder, 2001, p. 70).
 10. In preserving the frequency-time nature of sound configurations, ‘geometry must be the primary element in representation, primordial to other more abstract, symbolic and/or numerical representations. [...] Also, geometric representations of qualities as shapes have the advantage of being more in accordance with the approximate nature of perception and cognition, [...]. In this respect, cognition by shapes is well in accordance with connectionism principles [...].’ (Godøy, 1997, p. 94).
 11. It is still debated whether the component of the ‘Baddeley working memory model’ – the phonological loop – is specialized in processing only language or also other meaningful sounds, in particular music (Baddeley, 2012). According to Schulze and Koelsch (2012), the phonological loop can be involved in music processing or even incorporate subsystems specialized in processing non-phonological information (i.e., music). Moreover, Robert G. Crowder and Morton has elaborated on the idea of a ‘precategory acoustic storage’ for linguistic perception that is a system where acoustic material is subject to overwriting and to decay with time and where STM and LTM definitions lose their meaning and instead the presence of such a perceptual field would permit a series of straightforward conclusions to be made about immediate memory (1969).
 12. The phenomenal present of our consciousness develops over time in the form of an ever-moving ‘now’ and neurophysiology is currently investigating the experiential phenomena of timelessness which – in its most extreme cases – ‘can occur during near-death experiences, during intense suffering and emotions, violence and danger, altered states of consciousness, concentration and meditation, and shock’ (Fingelkurts et al., 2010, p. 216). The latter cases are often invoked in listening stances of certain practices within CAM (e.g., Oliveros’ deep listening; Voegelin, 2010; López, 2004).
 13. Linear temporality of traditional music (e.g., familiar or popular songs, well-known classical music) has been proved to variably match the perceived and imagined musical tempo (Jakubowski, 2020, p. 189).
 14. The idea of a perceptual spectrogram is a metaphor: it derives from the fact that our sensation of these Gestalt configurations appears to be very well described in terms of frequency/time/intensity (i.e., the dimensions of a spectrogram). However, several studies make use of neurograms to represent the incoming sounds in the auditory nerve (i.e., the firing rate distribution across this population of auditory nerve fibres). This ‘neurogram in many ways resembles the short-time spectrogram’ (Schnupp et al., 2011, p. 80).
 15. Dimensions are integral when it is not possible to assign an object a value in one dimension without giving it a value on the other. Three of these integral dimensions – measured as electric potentials for neuronal activity – comes directly from the first transduction of frequency, time, and dynamic.
 16. Godøy’s account, in particular, prefigures an epistemological foundation for this paradigm that consists in the idea that ‘musical objects [could be seen as] as a retentional image in memory’ (1997, p. 99).
 17. On this matter, we may find an important reinforcement in the work of Gallistel which shows evidence that spatial representations in the nervous system have a vectorial form (i.e., functioning like a vector space) (1990).
 18. In a similar vein, Nicod (1924) wonders what conception of the world could a being have with no sense other than hearing.
 19. Here, following Swanson, we will use the term ‘thinghood’, although in the recent literature on auditory objects (see, for a review, Griffith & Warren, 2004), the term ‘objecthood’ seems to be more appropriate. The concept of ‘auditory object’ has recently attracted much attention among

neuroscientists, neuropsychologists, and philosophers. Although an auditory object is a sense-specific concept, it 'raises fundamental questions about the nature of object perception in all sensory systems' (Griffiths & Warren, 2004, p. 887).

20. In audition, there are edges in pitch, and edges in time, but not in space. The claim that there are edges in pitch may seem strange, but a moment's thought will show that the idea is quite natural. A biologically important source is periodic, at least over short periods of time. Therefore, it is characterized by a fundamental frequency that can be thought of as its lower edge in pitch. The heard individual sound is further characterized by the shape of its leading edge in pitch-time, its attack, its trailing edge in pitch-time, and its decay (i.e., ADSR features).

References

- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). 'What' and 'where' in the human auditory system. *Proceedings of the National Academy of Sciences*, 98(21), 12301.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1–29.
- Bériachvili, G. (2010). *L'Espace musical: Concept et Phénomène à travers l'Avant-Garde des Années 1950–60 (Stockhausen, Xenakis, Ligeti...)* (PhD Thesis) Université de Rouen.
- Blauert, J. (1996). *Spatial hearing. The psychophysics of human sound localization*. MIT Press.
- Bonnet, F. J. (2012). *Les Mots et Les Sons: Un Archipel Sonore*. Éditions de l'éclat.
- Bourdieu, P. (1979). *La distinction. Critique sociale du jugement*. Ed. de Minuit.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.
- Chouvel, J.-M., & Solomos, M. (1998). *L'espace: Musique / Philosophie. Collection Musique et musicologie*. L'Harmattan.
- Clarke, E. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford University Press.
- Corbett, J. (2000). Experimental oriental: New music and other others. In G. Born & D. Hesmondhalgh (Eds.), *Western music and its others: Difference, representation, and appropriation in music* (pp. 163–186). University of California Press.
- Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 5(6), 365–73.
- Davies, S. (1994). *Musical meaning and expression*. Cornell University Press.
- Demers, J. (2010). *Listening through noise: The aesthetics of experimental electronic music*. Oxford University Press.
- Degerman, A., Rinne, T., Salmi, J., Salonen, O., & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. *Brain Research*, 1077(1), 123–134.
- Dibben, N. (1999). The perception of structural stability in atonal music: The influence of salience, stability, horizontal motion, pitch commonality, and dissonance. *Music Perception*, 16(3), 265–294.
- Egermann, H., & McAdams, S. (2013). Empathy and emotional contagion as a link between recognized and felt emotions in music listening. *Music Perception*, 31(2), 139–156.
- Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception*, 23(3), 221–247.
- Eitan, Z., Granot, R. Y., & Ornoy, E. (2012). Listening in the dark: Congenital and early blindness and cross-domain mappings in music. *Psychomusicology: Music, Mind, & Brain*, 22(1), 33–45.
- Evans, G. (1980). Things without the mind. In Z. Van Straaten (Ed.), *Philosophical subjects; Essays Presented to P. F. Strawson* (pp. 76–118). Clarendon Press.
- Farthofer, L. (2007). *Georg Friedrich Haas: Im Klang Denken*. PFAU.
- Fingelkurts, An. A., Fingelkurts, Al. A., Krause, C., Kaplan, A., Borisov, S., & Sams, M. (2003). Structural (operational) synchrony of EEG alpha activity during an auditory memory task. *NeuroImage*, 20, 529–542.
- Fingelkurts, An. A., Fingelkurts, Al. A., & Neves, C. F. H. (2010). Natural world physical, brain operational, and mind phenomenal space-time. *Physics of Life Reviews*, 7, 195–249.
- Frey, A., Daquet, A., Poitrenaud, S., Tijus, C., Fremiot, M., Formosa, M., Prod'homme, L., Mandelbrojt, J., Timsit-Berthier, M., Bootz, P., Hautbois, X., & Besson, M. (2009). Pertinence cognitive des unités sémiotiques temporelles. *Musicae Scientiae*, 13(2), 415–440.
- Furukawa, S., Washizawa, S., Ochi, A., & Kashino, M. (2013). How independent are the pitch and interaural-time-difference mechanisms that rely on temporal fine structure information?. In B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, & H. E. Gockel (Eds.), *Basic aspects of hearing: Physiology and perception* (pp. 91–99). Springer New York.
- Gallistel, C. R. (1990). *The organization of learning*. MIT Press.
- Gärdenfors, P. (2004). *Conceptual spaces. The geometry of thought*. MIT Press.
- Gibson, J. J. (1977). *The ecological approach to visual perception*. Houghton Mifflin.
- Giordano, B. L., & McAdams, S. (2010). Sound source mechanics and musical timbre perception: Evidence from previous studies. *Music Perception*, 28(2), 155–168.
- Glover, R., Gottschalk, J., & Harrison, B. (2018). *Being time. Case studies in musical temporality*. Bloomsbury.
- Godøy, R. I. (1997). Knowledge in music theory by shapes of musical objects and sound-producing actions. In M. Leman (Ed.), *Music, gestalt, and computing. Studies in cognitive and systematic musicology* (pp. 89–102). Springer-Verlag.
- Goodman, S. (2010). *Sonic warfare. Sound, affect, and the ecology of fear*. MIT Press.
- Grebosz-Haring, K., & Weichbold, M. (2020). Contemporary art music and its audiences: Age, gender, and social class profile. *Musicae Scientiae*, 24(1), 60–77.
- Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *The Journal of the Acoustical Society of America*, 61(5), 1270–1277.

- Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object?. *Nature Reviews Neuroscience*, 5(11), 887–892.
- Grimshaw, M., & Garner, T. (2015). *Sonic virtuality. Sound as emergent perception*. Oxford University Press.
- Grisey, G. (1987). Tempus ex machina: A composer's reflections on musical time. *Contemporary Music Review*, 2(1), 239–275.
- Hasegawa, R. (2015). Clashing harmonic systems in Haas's Blumenstück and in vain. *Music Theory Spectrum*, 37(2), 204–223.
- Hervé, J.-L. (2004). Formes et Temporalité dans les dernières Oeuvres de Gérard Grisey. In D. Cohen-Levinas (Ed.), *Le Temps de l'Écoute. Gérard Grisey ou la Beauté des Ombres Sonores* (pp. 16–21). L'Harmattan.
- Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge: MIT Press.
- Ikeda, R. (1996). Linear notes of 'Ryoji Ikeda +/-' (Touch, UK, ez-19017-CD).
- Imberty, M. (2005). *La Musique Creuse Le Temps. De Wagner à Boulez: Musique, Psychologie, Psychanalyse*. L'Harmattan.
- Jakubowski, K. (2020). Musical imagery. In A. Abraham (Ed.), *The Cambridge handbook of the imagination* (pp. 187–206). Cambridge University Press.
- Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press.
- Joseph, S., Kumar, S., Husain, M., & Griffiths, T. D. (2015). Auditory working memory for objects vs. Features. *Frontiers in Neuroscience*, 9, 13.
- Juslin, P. N., Liljeström, S., Västfjäll, D., & Lundqvist, L.-O. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–642). Oxford University Press.
- Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. Praeger.
- Kendall, G. S. (2014). The feeling blend: Feeling and emotion in electroacoustic art. *Organised Sound*, 19, 192–202.
- Kiss, J. (2004). *Composition Musicale et sciences cognitives*. L'Harmattan.
- Kivy, P. (2002). *Introduction to a philosophy of music*. Clarendon Press.
- Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *TRENDS in Cognitive Sciences*, 9(12), 578–584.
- Koffka, K. (1935/1999). *Principles of gestalt psychology*. Routledge.
- Kozak, M. (2020). *Enacting musical time. The bodily experience of new music*. Oxford University Press.
- Kramer, J. D. (1988). *The time of music: New meanings, new temporalities, new listening strategies*. Schirmer Books.
- Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organisation* (pp. 55–98). Erlbaum.
- Kubovy, M., & Schutz, M. (2010). Audio-visual objects. *Review of Philosophy and Psychology*, 1(1), 41–61.
- Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, 80, 97–126.
- Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. University of Chicago Press.
- Lehar, S. (2003). Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt bubble model. *Behavioral and Brain Sciences*, 26, 357–408.
- López, F. (2004). Profound listening and environmental sound matter. In C. Cox & D. Warner (Eds.), *Audio culture: Readings in modern music* (pp. 82–87). Bloomsbury.
- Massoud, M. R. (2017). *An analysis of Georg Friedrich Haas's in vain* (PhD Thesis). Northwestern University.
- McAdams, S. (1996). Audition: Cognitive psychology of music. In R. R. Llinás & P. S. Churchland (Eds.), *The mind-brain continuum: Sensory processes* (pp. 251–279). MIT Press.
- Mencke, I., Omigie, D., Wald-Fuhrmann, M., & Brattico, E. (2019). Atonal music: Can uncertainty lead to pleasure? *Frontiers in Neuroscience*, 12, 979.
- Menesson, C. (2008). Scelsi, le Temps ou la Respiration du Son. In P. A. Castanet (Ed.), *Giacinto Scelsi Aujourd'hui. Actes des journées européennes d'études musicales consacrées à Giacinto Scelsi (1905–1988)* (pp. 69–74). Publication Cdmc.
- Menger, P.-M. (2001). Il pubblico della musica contemporanea. In J.-J. Nattiez (Ed.), *Enciclopedia della Musica* (pp. 987–1004). Einaudi.
- Merleau-Ponty, M. (1945/1976). *Phénoménologie de la perception*. Gallimard.
- Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press.
- Nicod, J. (1924). Les relations des valeurs et les relations de sens en logique formelle. *Revue de métaphysique et de morale*, 31, 467–480.
- Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, 88(6), 1281–1296.
- Nudds, M. (2001). Experiencing the production of sounds. *European Journal of Philosophy*, 9(2), 210–229.
- O'Callaghan, C. (2010). Perceiving the locations of sounds. *Review of Philosophy and Psychology*, 1, 123–140.
- O'Callaghan, C. (2020). Auditory perception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2020 Edition). <https://plato.stanford.edu/archives/sum2020/entries/perception-auditory>.
- Ockelford, A., & Sergeant, D. (2012). Musical expectancy in atonal contexts: musicians' perception of 'antistructure'. *Psychology of Music*, 41, 139–174.
- Oliveros, P. (2005). *Deep listening: A composer's sound practice*. iUniverse.
- Oster, G. (1973). Auditory beats in the brain. *Scientific American, Inc.*, 229(4), 94–102.
- Raffman, D. (2003). Is twelve-tone music artistically defective? *Midwest Studies in Philosophy*, 27(1), 69–87.
- Reybrouck, M. (1997). Gestalt concepts and music: Limitations and possibilities. In M. Leman (Ed.), *Music, gestalt, and*

- computing. *Studies in cognitive and systematic musicology* (pp. 57–69). Springer-Verlag.
- Ridley, A. (2004). *The philosophy of music: Theme and variations*. Edinburgh University Press.
- Ross, A. (2010). Darkness audible. *The New Yorker*, 29 November 2010. Retrieved December 15, 2020, from <https://www.newyorker.com/magazine/2010/11/29/darkness-audible>.
- Santarcangelo, V., & Wanke, R. (2020). The early stage of perception of contemporary art music: A matter of time. *Organised Sound*, 25(2), 130–141.
- Schnupp, J., Nelken, I., & King, A. (2011). *Auditory neuroscience. Making sense of sound*. MIT Press.
- Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*, 1252, 229–236.
- Shamma, S. (2001). On the role of space and time in auditory processing. *TRENDS in Cognitive Sciences*, 5(8), 340–348.
- Sheets-Johnstone, M. (1999). *The primacy of movement*. John Benjamins.
- Snyder, B. (2001). *Music and memory*. MIT Press.
- Solomos, M. (2019). *From music to sound: The emergence of sound in 20th- and 21st-century music*. Routledge.
- Strawson, P. F. (1959). *Individuals. An essay in descriptive metaphysics*. Methuen.
- Swanson, J. W. (1967). On a problem of Nicod and Strawson. *Philosophy and Phenomenological Research*, 28(2), 222–229.
- Tenney, J. (1988). *Meta+Hodos: A phenomenology of 20th-century musical materials and an approach to the study of form; and, META Meta+Hodos*. Frog Peak Music.
- Tenney, J., & Polansky, L. (1980). Temporal gestalt perception in music. *Journal of Music Theory*, 24(2), 205–241.
- Terrien, P. (2012). *From composition to reception, identity of meaning and significance*. Proceedings of the Electroacoustic Music Studies Network Conference, EMS 2012: Meaning and Meaningfulness in Electroacoustic Music, Stockholm, Sweden.
- van der Cruys, S., & Wagemans, J. (2011). Gestalts as predictions – some reflections and an application to art. *Gestalt Theory*, 33(3/4), 325–344.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind*. MIT Press.
- Varèse, E., & Wen-chung, C. (1966). The liberation of sound. *Perspectives of New Music*, 5(1), 11–19.
- Varga, B. A. (2011). *Three questions for sixty-five composers*. University of Rochester Press.
- Voegelin, S. (2010). *Listening to noise and silence: Towards a philosophy of sound art*. Continuum Books.
- von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron*, 24, 95–104.
- Wanke, R. (2016). *Alvin Lucier. The space in the sound and the sound in the space. Linear notes of CD Alvin Lucier – Two Circles – Alter Ego*. Mode Record.
- Wanke, R. (2015). A cross-genre study of the (Ec)static perspective of today's music. *Organised Sound*, 20(3), 331–339.
- Wanke, R. (2019). A unified approach to perceiving sound in current experimental music genres. *Filigrane. Musique, esthétique, sciences, société*, 23. <http://revues.mshparisnord.org/filigrane/index.php?id=879>.
- Weale, R. (2005). *The intention/reception project: Investigating the relationship between composer intention and listener response in electroacoustic compositions* (PhD Thesis). De Montfort University.
- Wenzel, E. M. (1992). Localization in virtual acoustic displays. *Presence*, 1(1), 80–107.
- Windsor, L. W. (1995). *A perceptual approach to the description and analysis of acousmatic music* (PhD Thesis). City University.
- Windsor, L. W. (2000). Through and around the acousmatic: The interpretation of electroacoustic sounds. In S. Emmerson (Ed.), *Music, electronic media and culture* (pp. 7–35). Ashgate.
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47(1), 9–12.
- Zbikowski, L. M. (2008). Metaphor and music. In R. W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 502–524). Cambridge University Press.
- Zbikowski, L. M. (2017). Music, analogy, and metaphor. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 501–512). Routledge.