Musical Meaning and the Semiotic Hierarchy: Towards a Cognitive Semiotics of Music

Gabriele Giacosa

Abstract: Research on the meaning of music has a long tradition, with approaches from several fields, but it lacks a coherent framework for interdisciplinary discussions. As a result, the notion of meaning in music is fragmented among contrasting perspectives. I propose a cognitive-semiotic approach to the analysis of the meaning evoked by music listening, adopting a framework that eludes disciplinary limitations and expands the notion of meaning to the phenomenological concept of intentionality. For this purpose, I apply Zlatev's *Semiotic Hierarchy* to the experience of listening to music, analysing the diversity of meaning-making processes involved in music as distributed among several layers of experience. As a result, I propose an updated version of the Semiotic Hierarchy, clarifying its structure as based on possibilities of meaning-making, and allowing for temporality to pervade experience throughout all layers. I highlight the connectedness and simultaneity of different kinds of intentionality, resulting in the addition of the dimension of *aesthetic* experience – which I analyze as characterizing culture-general music listening.

A key claim is that experiencing music aesthetically articulates the listener's body in their inner sense of space and time, making them feel a sense of movement and vitality. This grounds music as a semiotic system, connecting with and fostering virtually uncountable subject-relative and culture-specific meaning-making acts.

Keywords: phenomenology of music, music listening, musical movement, vitality forms, intentionality

1. Introduction

Research on the meaning of music has a long tradition (Cross & Tolbert, 2016), with different perspectives coming from different fields, ranging from semiotics (Mazzola, 2003; Monelle, 1991) to cognitive neuroscience (Koelsch, 2011) and phenomenology (D. Clarke, 2011). Moreover, inter-disciplinary views have sprouted, connecting some of those approaches (e.g., Kühl, 2008), often in an evolutionary perspective (Patel, 2010), in comparison with language (Arbib, 2013). Nonetheless, research on music seems to lack a coherent framework for interdisciplinary discussions (Cross & Tolbert, 2020).

As a result, the notion of meaning in music is fragmented among contrasting perspectives (Cross & Tolbert, 2016; Almén & Pearsall, 2006). Traditional applications of semiotics, musicology and philosophy tend to focus on structure, with an emphasis on musical harmony, and remain enclosed in discipline-specific discussions (see Mazzola, 2003). On the other hand, empirical research (e.g., cognitive neuroscience and psychology of music) tends to move directly from pre-conceived views, reducing the experience of music to few culture-specific elements derived from theoretical analysis (Koelsch, 2011, 2019), while musical meaning remains under-explored (Cross & Tolbert, 2016).

The traditional comparison with language led to an unfruitful search for language-like semantics in music, resulting in conceptions of music as content-less (see Mazzola, 2002, p. 25); rather than expanding the notion of meaning beyond linguistic models, attempts are made to forcedly connect musical content as a *signified* (Mirigliano, 2011). However, a more general concept of *meaning* should go beyond such an approach, avoiding the tendency to reduce the complexity of meaning to a single dimension (Zlatev, 2018). Alongside the relevance of comparative studies of language and music (see Cross, 2011), we might learn more about language and other semiotic systems by considering their differences rather than their

similarities.[1] A proper semiotics of music will require a different approach, allowing for cross-field and trans-cultural discussions, while focusing on what constitutes the specificity of music.

In this paper, I adopt the approach of cognitive semiotics due to its ambitions towards "a unified theory of meaning" (Zlatev, 2002, 2009). In the process, I propose a cognitive semiotic framework imbued with phenomenology that broadens the notion of meaning, by developing Zlatev's (2018) *Semiotic Hierarchy*, and apply this to the experience of listening to music, describing the various layers of meaning-making that music evokes.

To reduce the influence of culture-specific musical meanings, I focus on the cross-cultural identification of music – that is, the experience of feeling "music" any listener could have with musical instances, even from different traditions. Thus, I shall avoid problematizations of the nature of music, relying on the assumption of the universal presence of music in human cultures (Mehr et al., 2019) and the innateness of the predisposition to develop the cognitive capacities for music (Malloch & Trevarthen, 2018; Seifert, 2020).

The structure of the paper is as follows. In Section 2, I present the theoretical background of this research, explaining the cognitive semiotic framework I adopt, the Semiotic Hierarchy model and my interpretation of it. In Section 3, I analyse the various layers of the model in relation to the kinds of intentionality involved and the acts of meaning-making when applied to music listening. Section 4 focuses on the inter-layer connections I identify in the analysis. Finally, in Section 5 I summarize my findings and suggest further lines of research that might originate from this paper.

2. Theoretical background

To give a proper account of meaning in music, we need analyses that allow for the inclusion of all facets of meaning involved in experiencing music. Subjective, interpersonal and cultural aspects of meaning should be taken as working simultaneously and interacting, unraveling both the specificity and the multifacetedness of musical meaning. The cognitive semiotic approach that I adopt allows to do this by avoiding reductionist tendencies.

Through the absorption of phenomenology, different kinds of meaning can be reconciled, conceiving of semantics (i.e., linguistic semantics as vehicle-content) as a form of *signification* (a relationship between *intentional* objects; see Section 3.3.3) – therefore, as a special kind of *semiosis* or meaning making. Sonesson (2012a, 2012b) argued that Peircean semiotics is consilient with Husserl's phenomenology. In fact, phenomenology is often implicitly applied in linguistics and semiotics when reflectively describing structures of meaning, and it is crucial in studying perception, focusing on the formal structures of consciousness (Sonesson, 2012a). Expanding views from Peircean semiotics to a broader phenomenological approach, cognitive semiotics can reach beyond some limitations of other disciplines, connecting the structures of subjective and intersubjective experience with quantitative measurements.

As a new transdisciplinary science of meaning, cognitive semiotics expands and connects the fields of semiotics, cognitive linguistics and cognitive science, with the help of phenomenology (Zlatev, 2015). Incorporating the phenomenological notion of *intentionality* (especially in Merleau-Ponty's approach) as the interconnectedness of subject and world, or openness-to-the-world, it is possible to expand meaning to the entire *lifeworld* (Zlatev, 2018). In such terms, meaning is the result of the co-constitution between "I" and the world, with the subject as locus of experience: instead of "subjects vs. objects," the world manifests as being-in-the-world. At the same time, since inter-subjectivity originates spontaneously from subjectivity (Zahavi, 2001), meaning is shareable with others, instead of being limited to one's mental

[1] A similar suggestion is made by Burling (2005, p. 16): "[w]e will understand more about the origins of language by considering the ways in which language differs from the cries and gestures of human and non-human primates than by looking for ways in which they are alike."

"representations." Such an approach allows trespassing conceptual separations between fields, with room for comparisons between, e.g., language, gesture and music (Zlatev, 2018).

On a more general level, cognitive semiotics also allows to overcome tensions between theoretical and empirical views through the employment of the *conceptual-empirical loop* (Zlatev, 2015) – a re-iterative feedback between experiential analysis and empirical studies. Such a loop begins with intuitions about specific concepts and views theories as an outcome of the process (Pielli & Zlatev, 2020), avoiding pre-conceived theories and empiricist reductions as well as moving directly from abstract theories to operationalized empirical hypotheses (and vice versa). Thus, instead of asking *what* music is, the focus is set on *how* it is manifested in one's experience – therefore, on musical meaning (or meaning-making). In line with the loop, my attempt towards a phenomenological cognitive semiotics of music begins with intuitions and analyses of the experience of music listening, corroborated through second-person ("empathetic") interpretations of relevant ideas within the literature.

According to the cognitive semiotic model known as the *Semiotic Hierarchy* (Zlatev 2009, 2018; Zlatev & Konderak 2022), experience is multi-layered, rather than the sole "aboutness" of thought. Meaning is viewed as a dynamic process (instead of static products) of *meaning-making*, distributed on all layers of experience. This is related to the temporality of experience and the possibility of shifting the focus of awareness, while including the pre-conditions for consciousness in corporeality (see Section 3.1). Layers of experience work simultaneously, in a somewhat hierarchical manner, through different kinds of intentionality. At the same time, each layer is subject to diachronic and enchronic processes of *motivation* and *sedimentation* – respectively resulting in novel-usage spontaneity and conventionalization (Devylder & Zlatev, 2020). Thus, within each layer the subject incorporates *norms* (in a very broad sense) that guide and support their actions and assessments in each situation (Zlatev, 2018, p. 4). The model is shown schematically in Figure 1.

Meaning level	**Kind of intentionality**	**Normative structures**	**Acts of meaning making**
Language	Symbolic intentionality	Symbols Syntax	Symbolic expression Linguistic expression
Sign function	Signitive intentionality	Signs	Sign use
Intersubjectivity	Shared intentionality	Empathy Conventions Communicative intent	Bodily communication Imitation
Subjectivity (pre-reflective self-consciousness)	Perceptual intentionality Inner time consciousness	Emotions The lived body	Feelings Actions Perceptions
Life/Animation	Operative and drive intentionality	Body schema Habits Affect	Movements Sensing

Figure. 1 The Semiotic Hierarchy, with higher and lower related levels related through the non-reductionist notion of *Fundierung*, and a dialectic relation between sedimented norms and spontaneous acts of meaning on each level, adapted from Zlatev (2018) and Zlatev & Konderak (2022)

The different *layers* of meaning-making are viewed as emerging upon one another under an irreducible originator-originated "foundational" (*Fundierung*) relationship, "where the lower level both provides the ground for the higher and is 'sublimated' by it:" the originator becomes manifest through the originated, which is a determinate form of the originator. Viewed as different "'planes of signification' or levels of meaning,"[2] characterized by different forms of intentionality (Zlatev, 2018, pp. 5-6), such layers are neither rigidly separate nor autonomous. Instead of drawing sharp distinctions, layers are connected, with higher ones consolidating the lower. Shifting from pre-reflective consciousness (life/animation) to reflective consciousness (focal subjectivity), to sharing and communication with others (intersubjectivity), in semiotic systems this continuity can work through a conjunction of elements: signitive intentions are indirect, mediated by signs, with a separation between the expression and the intended object – with language as a particular case based on symbol interplay.

However, Zlatev (2018) omitted to clarify a crucial aspect of the model. As I understand it, the Semiotic Hierarchy is a representational tool, meant to be neither static nor definitive. Layers of intentionality are *not* superimposed, separating aspects of experience or over-burdened with hierarchical implications. Rather, layers are open-ended and interact with each other. Moreover, the hierarchy does not represent cognitive *capacities* (in the psychological sense), but layers of *experience*. Switching between layers is a matter of directing one's focus to distinct aspects of experience, not of acquired "no-turning-back" stages: they are simultaneous possibilities of meaning-making.

There is, nonetheless, a partially hierarchical side to this description, which is related to phylo/onto-genetic acquisitions in terms of the possibilities of any subject's experience. In this sense, "higher" possibilities require "lower" ones. It is true that, phenomenologically speaking, drawing generalized assumptions about the experience of others is problematic. Yet, this is where phenomenological cognitive semiotics becomes crucial, supporting intuitions with behavioural and qualitative analyses (and vice versa). This allows to take some steps towards describing the "formal structures of experience," looking for the specificity of given forms of experience and their possible interactions.

Furthermore, Zlatev's model did not clearly display the potential simultaneity of acts of meaning-making between kinds of intentionality, and their relations. For instance, although signitive intentionality requires intersubjectivity and subjectivity, some sign usages might be simultaneous to acts in those layers – e.g., the emotions evoked when hearing an actor reading poetry are not "before" or "lower" than the experienced words. Moreover, acts in a lower layer might require acts from a higher one – as emotional responses evoked by watching a face depicted in a comic strip require the identification of the sign as a face and the attribution of expressivity to it. To overcome this limitation, I shall employ a different design for the model, highlighting some crucial inter-layer connections, as elaborated in Section 4.

Finally, specific acts of meaning-making could be interpreted as belonging to some other layers as well. In the analysis offered in the following section acts are located in the layer they are most prominently connected to, in accordance with the features considered in specific interpretations. Locating the same act in multiple layers would be pointless, as it would only imply focusing on different features that were interpreted as belonging to other layers or acts. Again, the model does not imply any temporal relationships between specific acts – which might happen simultaneously.

[2] Zlatev (2018) employs the terms *level* and *layer* interchangeably. I prefer the latter, similarly to Zlatev and Konderak (2022), as it suggests less-rigid borders.

3. The Semiotic Hierarchy and music

In this lengthy section, I apply and develop the Semiotic Hierarchy model presented in the previous section. As the focus of this research is to identify the specificity of musical meaning-making across its experiential "formal structures" (Thompson, 2007, p. 28), I shall analyse musical meaning from a general perspective. Culture-specific meanings will only be considered in general terms.

Applied to music, this model allows us to maintain all dimensions of the listener's experience of a piece, without reducing, for instance, subjective interpretations to a pre-conceived list of possible meanings. Rather, affect and emotions can fit the hierarchy as generally present in listeners' experiences, without a need for all listeners to experience the same feelings at all instances. Moreover, since layers are inter-connected, subjective meanings can be influenced by higher layers; thus, intersubjective sharing (and context), degrees of conventionality and cultures can influence individual feelings.

In what follows, I describe each layer in relation to the experience of music listening, highlighting some of the acts of meaning-making involved.

3.1. Life/animation: operative/motor intentionality

As the basic layer of meaning, we encounter *living/being* as openness-to-the-world, where individuals are guided by, in general terms, life-supporting and life-enhancing values. Despite seeming mostly inaccessible to phenomenological investigations, this layer constitutes the primary ground of interaction between a subject and the surrounding world, resulting in "pre-conscious"[3] actions and evaluations. This perspective, foreshadowed by Husserl's *operative* intentionality (*fungierende Intentionalität*) – the "antepredicative unity of the world and of our life" (Merleau-Ponty, 1962, p. xx) – was consolidated by Merleau-Ponty's notion of *motor* intentionality. Here, the subject is present as a *living body*, prior to objectifications of experience – that is, this kind of intentionality is not a matter of "aboutness," but rather of the pure connection between the body and the surrounding world.

The living body self-organizes through what Merleau-Ponty calls *body schema*, a pre-conscious system of possibilities and minimal evaluations, governing movements and (re-)actions and acting as the ground for sensing – the minimal core of the self. Merleau-Ponty (1962) highlights how *motility* (the potential to move) constitutes this ground in terms of potential interactions of "I can" rather than "I think that." The body, in its potential for movement and action, is not a mechanism: it is the "vehicle of being in the world" (Merleau-Ponty, 1962, pp. 158-160), grounding consciousness in the minimal sense of agency. Normatively, this is the realm of habits and affect, as related to the continuous relevance for the primordial self of movements and sensing (Zlatev, 2018, p. 7).

3.1.1. Husserl's analysis of time-consciousness

Due to the temporal nature of music, phenomenological views on music often start by considering Husserl's description of the experience of time (Husserl, 1991; see D. Clarke, 2011; Montague, 2011), where he analyzed it in relation to a listener's experience of the sounds composing a musical melody. Nonetheless, most of the musicological readings of Husserl's description seem to misunderstand his attempt as a description *of* experiencing melodies, whereas his intent was to describe the dimensions in which time is experienced through the example of a melody.

[3] *Pre*-conscious is used instead of "un-conscious" (Zlatev 2018, p. 9) to highlight the connectedness between layers of awareness: it is not a biological process evading awareness, but rather the core of awareness.

In order to clarify this, we need to assess the relationship between the layers of operative and reflective intentionality, especially in relation to time. Despite them being substantially continuous, it is possible to sketch a line between the two. Husserl's (1991) description of time-consciousness is split in two dimensions: *the continuum (*or flux) and the *pure now*. The latter constitutes a "primary sensation," what is happening in the moment, whereas the former is manifest in awareness through a "primary memory" or *retention* of the events occurred in the preceding now and the *protention* of what is just to come. Yet, Husserl describes the *now* as an abstraction, an "ideal limit" that "we can only actually experience […] through what is other than it, through its retention" (D. Clarke, 2011, p. 5), while he also asserts that time consciousness is a primary necessity for all consciousness. Furthermore, *retention* is distinguished from *recollection* – the active recalling of (and potential reflection upon) memories. Thus, the now is the (extended) present of experience, but it happens before any one act of reflective consciousness (Zahavi, 2003).

As mentioned, Husserl also considers the future-oriented counterparts, respectively *protention* and anticipation/*expectation*, although they received less attention due to the focus on past and memory. I chose not to discuss these notions in detail here, because there is less literature and a proper analysis exceeds the scope of this paper. Although future-orientation is clearly involved in the constitution of moments and of musical experiences, I believe it is mainly related to the tensions involved in the emergence of experience: it would probably do little towards clarifying the experiential structures of music listening. Phenomenological analyses of further aspects of musical experience (e.g., timing, culture-specific meanings, improvisation) should focus on it properly.

It is crucial to remember that Husserl is describing the "formal structures of experience" (Sokolowski, 2008), the necessary pre-conditions for consciousness. Experience appears in time consciousness, which is constituted by a sequence of ideal points in the flowing of time. The pure now is precisely a logical pre-condition that one cannot experience as such. The acts pertaining to motor intentionality are intrinsically temporal: the experientable now *extends* (if minimally) through retention into the immediate past and protention into the immediate future. On the other hand, the logical (quantifiable) analysis of time is manifest only through recollection.

This creates a discrepancy between the (operative) act of sensing/perceiving and (reflective) focal awareness on the higher layer in the hierarchy, although their connection is so habitual that it is hard to imagine them separately. No temporal or ontological implication is present in the model: one does not need to recall a whole melody to perceive the notes that constitute it, as the now is "before" the flux (of objectual time) connecting all points in time. On the other hand, as subjects we do not have focal access to our operative intentionality, therefore we only become aware of that layer *a posteriori* – in a similar manner to one realizing they were eating their nails unwillingly in a stressful situation.

Finally, we can turn back to Husserl's melody example. According to him, sound events appear in a listener's mind on the background (*horizon*) of silence; each event establishes a now-point loaded with the retention of previous events (and *protention* for further events). The time-flux is the last component to emerge (Montague, 2011, pp. 34-36), as it is experienced in recollection. Accordingly, Husserl studies melody *qua* temporal object, assuming it as a temporal unit (Husserl, 1991, pp. 22-23): we are pre-reflectively aware of sounds in retentions and protentions, but reflectively aware of melodic units through recollection *a posteriori*.

Surprisingly, Gallagher (1998) criticized Husserl's description of the "linearity" of melodies "for ignoring the diverse temporal possibilities in the act of listening" (see Montague, 2011, p. 35). As Merleau-Ponty had pointed out, listening to a melody requires such linearity (Merleau-Ponty, 1962, p. 474). Any kind of listening that is not linear implies a further conscious control over the material dimension of listening – listening in a specific way, as focusing on specific aspects or features – thus requiring recollection and memory (commonly interpreted). Control over forms of listening thus properly belongs to the layer of reflective intentionality, discussed in Section 3.2.

Similarly, Montague controversially employs Husserl's description to analyze the experience of playing (and listening to) musical pieces. In particular, he speaks of "a network of retentions and protentions as defined through bodily gesture" (Montague, 2011, p. 39) in relation to the recollection that emerges when playing recurring melodic bits in a piece, thus blending the pre-reflective and reflective sides of temporal experience through the identification of similarities and repetitions of melodic (and gestural) units. Larrabee suggests that the identity/similarity of various events – even in direct succession, as "the smelling of the same bread" (Larrabee, 1989, p. 380) – implies a *recognition* of events, therefore recollection. It seems safer to limit Husserl's description of time consciousness to pre-reflective components, as the meanings related to the recognition of structural order in musical pieces are on a different scale from those related to pre-reflective experience.

In conclusion, although minimal tone sequences are experienced in retention (and protention), the precise temporal and pitch relations required for experiencing proper melodies pertain to reflective intentionality, and thus a higher layer of meaning-making. On the other hand, operative/motor intentionality includes acts of meaning-making that are related to the experienced (extended) now. This excludes pitch relations, melodies, and temporal structures, while allowing for the features of single tones – such as (pure) pitch, timbre, and loudness – and minimal sequences.

Following Montague, it is true that, at least for musicians, the gestural aspect of playing music involves a form of corporeal *habit*, pertaining (if partially) to operative intentionality – exemplified by the subtle uncontrolled movements of the right hand required to play an *arpeggio*, although the guitarist's attention is on their left-hand fingers. Nonetheless, these acts only turn into habits through practice and expertise; moreover, some of them are specific to playing music rather than listening. Limiting my view to what a non-musician listener might experience of the corporeality of musical sound in this layer, we need to distinguish the features that derive from a mimetic act of *intersubjectivity* with the musician – thus pertaining to a different layer (see Section 3.3) – from those related to the music *per se*.[4] Thus, whereas perceiving a musician's bodily movements and interpretation features (such as *vibrato*) is mainly an intersubjective act, we can include here features related to the sound-producing actions that determine the "grain" (Godøy, 2011, p. 232) of the sound – e.g., *attack* and *decay* stages in envelope synthesis.[5] In the following sub-section I explore these acts of meaning-making and the values linked to them.

3.1.2. Movement, affect and sub-chunking

In *The Visible and the Invisible* (1964/1968) Merleau-Ponty criticized an aspect of Husserl's phenomenology of time, suggesting that Husserl failed to consider the "time of the body" and "of the corporeal schema:" the ideal now-present should not be viewed as immanent, but as transcendent; it is a "symbolic matrix and not only *a* present that breaks up toward the past" (Merleau-Ponty, 1964/1968, pp. 173, 192; emphasis in the original). The relation between the

[4] In this case, different layers interact to a high degree, making it hard to draw a line. As clarified in Section 3.3, I consider music to be experienced as expressive in its own regard, thus attributing corporeality to it through empathy and motor intentionality, whereas features linked to another subject's experience (here, movement) seem more related to intersubjectivity, as the absence of the other subject would prevent their emergence.

[5] These are the first two stages typically identified/employed in envelope generators in sound synthesis. *Attack* refers to the time required for the sound volume to rise from nil to peak, while *decay* is the time required from peak to *sustain* (the volume maintained for the duration of the sound). Together with the timbre of the instrument (its "color," how the materials employed resonate in terms of frequencies), attack and decay (and partially sustain) are crucial to describe how a sound "appears:" the sweet immediacy of a piano, the slower grittiness of a violin, …

temporal now and the body schema reveals the radical connection between intentionality and meaning-making: even the basic layer of life implies a value system with its normative structures.[6]

In my initial description, I had identified habits as the typical acts connected to the layer of motor intentionality. In line with this and with Carman's example, *movements* emerge as fundamental acts of meaning-making, grounding the relationship between self and world. In this regard, it is possible to make a distinction between inner *movement* as the self-rooted corporeal sensing of dynamism/vitality (from self or other entities), and outer *motion* as the thematic/reflective perception of moving objects in the surroundings (Sheer-Johnstone, 2011).

Merleau-Ponty develops an analogous dichotomy in relation to the body and movement in space (and time), opposing "orientated space" (living, pre-reflective) to "homogeneous space" (reflective) (Merleau-Ponty, 1962, p. 117). His dichotomy is rooted in the double sensations of the body inhabiting a space/time. Self-movement can be experienced as both inner and outer (see Sheets-Johnstone, 2011, pp. 514-518), depending on a shift in focus. This shift, though, highlights the connection between pre-reflective perception and reflective perception: in the layer of life, movements (of the self or from external objects) are experienced as *dynamic*, but not in relation to a geometrical space – which requires recollection and, thus, reflection. For instance, one does not typically focus on the experience of running when trying to catch a bus. Yet, when noticing a stone on the ground or sensing a "wrong step," they will focus on that experience; furthermore, while training for a marathon they will likely reflect on various details/moments of the motion – in order to improve their performance.

Additionally, Sheets-Johnstone (2012) identifies the inner sense of movement as the source of experience of both external motion and emotions (but see Zlatev, 2012, p. 7). Referring to Stern's (1985) notion of "vitality affects" as bodily-kinetic dynamics, she describes movements as generating and *articulating* dynamic-affective meaning through the body (Sheets-Johnstone, 2012, pp. 31-34). In other words, the living body self-organizes through movements and postures, discretizing time and space, in relation to the values of dynamic self-world interactions. This suggests a distinction between *affect* and *emotion* – the latter being related to the reflective identification of patterns of affective dynamics. Stern describes "everyday" emotions as composed of a "vitality form" (dynamic affect) and a discrete emotion (learnt blend of quality and tendencies, e.g., *joy* as positive and distinct from other states) (Stern, 2010, pp. 27-28). Thus, despite their connection, I refer affect to the layer of life/animation, and emotions to reflective consciousness.

Let us now focus on music, and to the qualities and effects of sounds in a listener's experience. Despite the difficulty of isolating pre-reflective and reflective consciousness, a comparison with cognitive psychology might be fruitful – with no intent of directly connecting psychology and phenomenology. An interesting attempt in this direction was taken by Godøy (2011), who focused on the connection in music awareness between sounds and actions, for producing sounds or other related actions. Most of the results are related to a simulation of movements, which must involve a degree of reflective consciousness. Yet, focusing on the temporal spans that shape the experience of music in time – and through a comparison with Husserl's analysis of time consciousness – he identifies "three different concurrent timescales in the perception and cognition of music" (pp. 240-241): *chunk*, *sub*-chunk, and *supra*-chunk. Whereas chunk and supra-chunk comprise temporal spans related to recollection, the sub-chunk is assumed as a basis for the time-awareness segmentation ("chunking") process and involves minimal temporal spans. This strongly resonates with the notion of operative intentionality and the now: Godøy's sub-chunk level is conceived of as "continuous sound," and it involves the features of "pitch, timbre, and loudness" (p. 241) – supporting my suggestions from the previous sub-section. The value of these features might be related to material adjustments to the surroundings, such as the proximity of a moving object or the potential danger of an entity. Think

[6] Zlatev (2018) follows Carman in identifying normativity for this layer in the "felt rightness and wrongness of the different postures and positions we unthinkingly assume and adjust throughout our waking (and sleeping) lives" (Carman, 2005, as cited in Zlatev, 2018, p. 4).

of the different effects of hearing a fast metallic sound or the soft sound of something falling into a liquid.

3.2. Focal subjectivity: reflective intentionality

In fact, *passe* Zlatev (2018), subjectivity as pre-reflective (but "reflexive") intentionality already belongs to the first layer, as examples given above have shown. However, as repeatedly noted, the line to *reflective* (or focal) intentionality is not thick: self-conscious subjectivity is founded (the *Fundierung* relationship) upon and emerges from the "minimal experiential self" (Zlatev, 2018, p. 8) of pre-conscious awareness. Reflective awareness turns to itself, allowing for "streaming consciousness" to become "a thematic noticing" in itself (Husserl, 2001, p. 320, as cited in Zlatev, 2018). Husserl's "primary memory," retention, becomes the foundation of reflection, allowing for the emergence of the temporal structures of *recollection* and *expectation*. Events are identified as distinct – and as separate from the self.

This thematic noticing highlights a crucial aspect of awareness that was not accessible in operative intentionality: the potential for a subject to shift (in part volitionally) their attention to different features of experience. While retention is a somewhat involuntary process, the ability to focus attention allows for a kind of control, so that one can compare different memories in time and make expectations about specific objects or events – to the point of reflecting upon reflecting.

Focal subjectivity is the realm of self-conscious perception, including being aware of one's own interaction with the world. Guided actions result in feedback evaluation, which can be sedimented to further foster control. On the other hand, control over behavior and events, mixed with the possibility to reflect upon evaluations (both towards the past and the future), leads to the emergence of specific qualities of feelings and *emotions*, as mentioned earlier and elaborated below.

3.2.1. Time structures and chunks

To describe this layer in relation to music listening, I first need to assess the constitutive role played by temporal structures. In fact, on the one hand reflection allows for thematization, the "aboutness" of experience, so that "sounds are experienced as the sounds of things" (Ihde, 2007, p. 85). On the other hand, reflection dwells upon and is intertwined with retention: both have their own temporality and appear within the flux of time. Sounds are not static entities appearing just in the now: a sound is always intrinsically related to motion and the passing of time. Since the subject is always imbued in temporality, experience of any event/object is temporal. Nonetheless, sounds have a peculiar way of highlighting the temporality of experience, as their structures and *gestalts* derive mainly from their duration, passing, and relations, and in a way "sound embodies the sense of time" (Ihde, 2007, pp. 83-85).

Ihde (2007) shows how the attentive *temporal span* can be modified by shifting one's auditory focus to broader or smaller elements (pp. 86-90). For instance, in a conversation one does not normally listen to single sounds or syllables, but rather to larger bits of flowing sounds; afterwards, one can recall and isolate single elements. Similarly, one can predispose themselves to attentively listen "for" specific features *a priori*, as an expectation – as when waiting for the starting signal of a running race, where the passing by and succession of sounds becomes less relevant in awareness. This implies that different ways of listening can be employed; here we find room for Gallagher's criticism of Husserl's "linearity" of listening. Still, rather than contrasting one another, different modalities tend to interact: while focusing on remembering previous passages from a musical piece, one does not cease to experience the hearing of ever-new retentions. As a result, several temporal structures can be experienced with music: one experiences the flowing linearity of sounds in time while also constructing rhythmic/melodic bits and an overall structure of the elements – or thinking back to previous (or future) elements.

At the same time, in thematic perception the subject identifies objects and, through recollection, can establish connections between their features. Such modality is sedimented into our natural attitude towards the world, to the degree that focusing on sounds alone is an exception – as demonstrated by Schaeffer's work on *musique concrète* (Schaeffer, 1966; see Godøy, 2011; Schiavio et al., 2017).

This helps clarify our previous description of the backwards "influence" of recollection on retention: our habitual experience is so grounded in reflective awareness that, when exposed to sounds, we are immediately drawn to identifying their sources and features – or perhaps following their features and further contents (as in language, through interaction with other layers of meaning). As with smelling "the same bread" (Larrabee, 1989, p. 380), recognizing the same sound sources or melodic sections is an act of recollection – albeit seemingly immediate; retentional structures lie beneath the surface of recollection and require a specific act to be actively noticed. This is an implicit (yet crucial) process in listening to music, as melodic lines are typically constructed through the identification and connection of the sounds produced by separate instruments (or groups of similar ones, as with strings sections), even within complex polyphonic interactions between melodies.[7]

Ihde's view on temporal spans and auditory focus mostly overlaps Godøy's concept of chunking (see 3.1.2), based on distinct temporal spans. Considering the discussed continuity between retention and recollection, whereas Godøy's sub-chunk level was conceptually close to the now, his separation between *chunk* – a span "typically in the 0.5–5 second range" – and *supra-chunk* – a "concatenation of several chunks" (Godøy, 2011, p. 240) – runs the danger of being too strictly time-dependent,[8] not highlighting the reciprocal influence between temporal structures. Nonetheless, we can partially refer to his model to identify the acts of meaning-making involved in the layer of focal subjectivity.

Without drawing clear lines, I suggest that with minimal focal attention a listener[9] experiences minimal temporal *sequences*, pitch contour and relations, and harmonic interactions (supposedly in an implicit form of dissonant/consonant). Through recollection, then, time and pitch are discretized into more and more clear structures, creating properly *rhythmic* and *melodic* sequences (as in a march *vis-à-vis* a shuffle bit, or identifying the first few notes of a song). Further acts of reflection, even in very short temporal spans, lead to the identification of broader structures, such as rhythmic regularities (as 4/4 or 6/8), motifs, and explicit harmonic relations (e.g., chords). As acts of recollection can build upon one another, in larger time spans a listener can identify specific pieces, internal relations (as a *leitmotif*), and even recognize similarities to other pieces.

This description is not in contrast with Husserl's view of melody as a temporal unit, as time consciousness permeates and substantiates all the higher layers, albeit in different ways. Melodies may well be conceived as the temporal units that delineate retentional structures, but they do so backwards: a melodic bit is not perceived as such until recollection is acted upon it – since we are not fully aware of this process until it is ruled out. We could relate this to Merleau-Ponty's description of *gestalts*:

> The *Gestalt* is not a spatio-temporal individual, it is ready to integrate itself into a constellation that spans space and time—but it is not free in regard to space and time, it is not aspatial, atemporal, it only escapes the time and space conceived as a series of events in themselves. (Merleau-Ponty 1964/1968, p. 205)

[7] Ihde (2007) describes various ways of listening, but that exceeds the scope of this paper. I considered some of the aspects that help analyze the typical processes involved in musical meaning-making. Suffice it to say that applying a specific listening "stance" implies reflection and attentional control.

[8] Specifically, 5 seconds seem a far too long span for *pre*-reflective awareness. Godøy's model seems too close to the psychological approach (rather than phenomenological): temporally, retention and recollection can overlap – and it is possible that 0.5s is enough for *some* reflection to happen.

[9] I consider the case of a non-musician listener. It might be the case that a trained musician develops a non-reflective ability to identify some features more clearly.

As a final clarification, I wish to emphasize that the acts here described do not define music specifically: these are acts of meaning-making involved when listening to music, especially regarding sounds, but they mainly concern *listening*.

3.2.2. Emotions, *Vorstellung* and mind-wandering

The connection between operative and reflective intentionality is also evident when considering the normative structures and values involved in focal subjectivity. In Section 3.1.2, I referred to Stern's distinction between affect and emotions. Accordingly, although they are experienced in connection, emotions emerge in reflective awareness because they imply the identification of discrete qualities/features; on the other hand, vitality affects are "a-specific" and motivate/ground the experience of emotions (Stern, 2010, pp. 27-28).

With regards to music, the first components to consider here are the effects related to the temporal structures previously described – that is, the results of hearing a minimal sequence or a screeching sound associated with a potential danger, as opposed to a soothing sound. A similar view is intrinsic to the notions of harmonic dissonance and consonance, which derive from physical features of the sounds. Moreover, a history of associated emotions can sediment onto specific sounds in a subject's memory – for instance, the horn of a truck could evoke a particular fear in someone who experienced a car-truck accident.

The elements described so far only consider the layers of operative intentionality and focal subjectivity. Nonetheless, through interaction with further layers of meaning, further subjective experiences can emerge. The acquisition of systematized musical signals (in intersubjectivity, see 3.3.3) often gives rise to variable subjective responses, from specific emotional responses to melody-concept associations – e.g., when a listener freely interprets (or develops an experiential association) a melody as descriptive of specific events in their life. This kind of association pertains to focal subjectivity because it is not a matter of *shared* experience in communication with others. In this sense, it resembles Frege's notion of an individual's "conception" (*Vorstellung*) as opposed to the sense of an expression or word (*Sinn*) – a position interestingly held by Husserl as well (see Hilpinen, 2015, pp. 992-993, 1008-1009). A further experience derived from such associations (if partially) is that of *mind-wandering*, *i.e.,* when a listener finds themselves "drifting away" in their own subjective associations, partially detaching from the music.[10]

3.3. Intersubjectivity: shared intentionality

The relation between subject and world implies the possibility of relating to other subjects, leading to the layer of *shared* intentionality. The interplay of subjects is a complex and multifaceted experience, but it is possible to identify some of its internal structures.

Intersubjectivity is often analyzed in terms of projections and simulations of one's experiences "into" others. Phenomenology regards this as a specific form of intersubjectivity, definable as *sympathy*, while including in this layer a more basic form termed *empathy* (*Einfühlung*) – "the awareness of another's experience" (Thompson, 2007, p. 386), related to a specific kind of intentionality. Through empathy we experience others as intentional entities, able to express their own awareness. The most basic form of empathy – "sensual empathy" or "sensing-in" (Thompson, 2007, p. 389) – emerges from the duality of the body as subject (*Leib*) and object (*Körper*), without this implying any ontological dichotomy (see Zahavi, 2001). At this level we find a *direct* perception of the "bodily manifestations" of the subjectivity of others

[10] This might involve a particular way of listening to music, but it does not detach from experiencing music *as such* – rather, it might be a form of meaning rarely considered by experts and musicologists, since it involves a natural attitude that does not focus much on musicological material.

through their corporeality and movements, perceiving "vitality forms" from others, with no "need to 'infer' or 'simulate'" their subjectivity (Zlatev, 2018, p. 10).

Empathy involves a substantially different viewpoint from the first-person perspective of the self, leading to a second-person perspective. Yet, first- and second-person perspectives are deeply connected. On the one hand, pre-reflective and reflective subjectivity (the first two layers) are a prerequisite and the foundational ground (*Fundierung*) for intersubjectivity. On the other hand, only through the relation with the *other* is the self established as an interpersonal *self* – in other words, I become aware of being specifically *I* by identifying the other as *other*. Moreover, this relation is the source of the notion of *objectivity*, thanks to "the awareness of others being oriented in perception and practical action towards the same objects as us" (Zlatev, 2018, p. 10).

Building upon these features, we find the domain of (intentional) *communication*, which requires awareness of a *separation* between the subjects, something not fully possible in the previous layers. Through basic empathy we can directly communicate (or fail to do so) with others, perceiving their subjectivity through their corporeality and actions: the "body expresses the perceptual possibilities of continuous synthesis and *intercorporeal* interaction" (Guareschi, 2019, p. 52, my emphasis). Then, a further level of communication is achieved via the employment of *communicative intent* – conceivable as a second-order intent to direct the other (both physically and mentally) and/or to have one's own expressive goals recognized. Finally, after the acquisition of the skills necessary for *bodily mimesis* (Donald, 1991), even more complex forms of communication emerge that imply growing degrees of volitional control to imitate others' actions. The sedimentation of such communication leads to social conventions and the foundation of *culture* – meant "as both practical (implicit) and theoretical (explicit) knowledge" (Zlatev 2018, p. 11).

3.3.1. Empathy in music

To analyze intersubjectivity in music listening, a key step is to assess the ways in which the structures of empathy manifest. Recently proposed analyses of empathy in music (van der Schyff & Krueger, 2019; Clarke, 2019) show how empathy is crucial to the activities and cognitive processes involved in music. Accordingly, music *scaffolds* (i.e., supports, elicits and shapes) individual and social cognitive capacities (Krueger, 2019), and it is used as a tool to construct "empathic spaces," and to explore and manipulate social and emotional contexts. Nonetheless, the focus of these studies is on the cognitive capacities involved and/or what we "do with" music in interactions with others, but less on describing empathy in the experience of music listening. My goal here is to describe the structure of empathy in relation to music listening.

At the most basic layer, the pre-reflective self perceives the other (presumably, a musician) as manifesting through sounds, as sounds afford otherness due to the corporeality of experience (*living body*), evoking Stern's vitality affects/forms. This is not a sympathetic projection ("like me") of my experience onto the other, but a direct connection lying at the core of the self: through empathy, the meaning-making that flourishes within motor intentionality can also *express* meaning in/from others, so that other bodies can evoke the same process in "my" body (see Thompson, 2001).

As reflective awareness emerges in the second layer, identifying the temporal structures of sounds, the self can experience (noetically) sounds as related to external sources (objects/events) and as potentially experientable from other subjects. Furthermore, an actual other person can be found to be the source of the sounds, establishing communication and sympathy, where simulations and inferential projections about the other's experience appear – such as the musician's *emotions*. Finally, through the sedimentation of communicative norms, specific forms of communication and their features emerge, resulting in systems such as music.

Only through this layered structure can a subject experience music in its communicative features, but as pointed out, these steps are not temporally consequential, but merely "structural." A listener does not always need to attentively identify another person as the source of sounds to

experience music – for instance, in some electronic music it can be hard to think of certain sounds as human-produced. Rather, in the words of a listener's first-person report, while listening "I [am] inside the music, and the music [is] inside me" (Gabrielsson, 2011, p. 86): it feels as a non-mediated connection. Nonetheless, this specific form of experiencing sounds would not be available without the gradual sedimentation of all the steps into a subject's experience (here, also in the sense of memory). All layers manifest simultaneously, while one's attention is focused on specific aspects from time to time – also implying that, when listening to music, we can switch back to focusing on the communication with the person producing the sounds in live contexts, or to analyzing the temporal/acoustic features of sounds.

The structure of empathy is also mirrored in the ways music is employed and experienced, through interaction with the lower layers in the hierarchy. On the one hand, music "acts as a 'technology of the self' that affords listeners the opportunity to structure and organize their subjectivities" (Clarke, 2019, p. 78): we can employ music to shape our corporeal (and emotional) experience in space and time. Accordingly, music is often conceived of in relation to "order." Yet, I suggest that the "shaping" and organization involved is different from that of external objectification (as with geometry, or the pentagram). It is a non-thematic corporeality rooted in the living body – as in Sheets-Johnstone's dichotomy of the feeling of *lived movement* as opposed to observed *motion,* see Section 3.1.2.

On the other hand, in Clarke's (2019) words, music can be used as "a medium for […] engagement with others, and an environment in which to explore" interpersonal relations (p. 79). For instance, considering the ability to synchronize rhythmically ("entrainment") with others, we first learn to tap together, sharing that experience, and then exploit the ability in musical contexts, until tapping together to music becomes a social tool that can enhance "empathizing" (p. 76). In this way, music allows us to articulate, express and share emotions interpersonally – a process thoroughly studied and explained in numerous ways (see Miu & Vuoskoski, 2017). The motoric layer grounds engagement between explicitly active participants (musician-audience, auditors dancing/clapping together), while reflectively one can project/infer the emotional engagement of other non-active auditors.

Before moving to communication, I shall focus on the experience of musical sounds as *moving*, highlighting how this constitutes the uniqueness of musical experience as aesthetic, grounded in empathy.

3.3.2. The aesthetic dimension: moving sounds

As mentioned earlier, Husserl (1991) conceived of melodies as *unitary* temporal objects (p. 40). Although it was not his actual focus, this conception suggests that musical structures afford an attribution of coherence and object-ness that surpasses "normal" listening. Whereas sounds are typically identified as objects in the world, musical sounds are felt as in focus *per se*. Nonetheless, we need to consider a further feature that does not emerge from Husserl's analysis: the feeling of musical sounds as *moving*.

Ihde (2007) begins his description of the modality of listening that is specific to music with the traditional notion that in music "sound draws attention to itself" (p. 155). He suggests that the "significance" of musical sound is not a "reference [to] things," but it "enlivens one's own body" to participate "in the movement of the music" (pp. 155-156). I propose that such experience of movement is central to music specifically. Ihde shows that other forms of acoustic manipulation/composition, attempting to evoke "a listening to the music of the World" (such as *sound art*), manage to draw attention to the sound as such, but in the end do "away with the idea of music" (p. 159). What is lacking in other forms of listening is precisely the emergence of melodies (and rhythms): musical temporal structures acquire a sense of unity and direction related to movement that involves a corporeality – both *in* the listener and *in* the sound.

Diana Deutsch's *speech-to-song* illusory perception (Deutsch et al., 2011) shows that, due to mere repetition, one can experience music through the same sounds and acoustic properties of

speech. Yet, listening to speech does not (usually) involve a sense of movement: one focuses on the cues that characterize communication (aboutness, emotion, prosody, etc.). With music, I am primarily drawn to the temporal flow of sounds, waiting for the next to start, while remembering the past ones: I focus on their dynamic relations and the features that constitute their forms of vitality (time, space, force, direction); I move/articulate my body aligning with these forms (beat/pulse, rhythm, melodic curve), *following* the sound.

Theoretical studies (Clarke, 2017; Teie, 2016) and empirical findings (Sievers et al., 2013) support the idea that movement is universally associated with music. Several studies report that music can even be experienced as a "virtual person" (Cochrane, 2010; Leman, 2007; Levinson, 2006; Livingstone & Thompson, 2009; Watt & Ash, 1998) with whom listeners engage empathically; listeners can personify "music itself as providing empathy and understanding," experiencing it "as a surrogate for an empathic friend" (Clarke, 2019, p. 77). For instance, when I listen to a nostalgic piano piece during a trip, staring at the landscape; or when, after achieving a goal, I listen to Tchaikovsky's *Violin Concerto*: I feel that music shares and understands my feelings, I am expressed in the connection with sounds. This *can* also involve a projection onto the musician/composer behind the music, but not necessarily: they are two distinct elements that can interact. Projecting onto the musician, I would only feel them as expressing themselves, with me being the "understanding one." Instead, when experiencing movement together with the moving sounds, I can also feel as expressing myself "towards" music.[11]

Both perception of movement and personification of the music imply the attribution of a vitality/subjectivity to the musical sound, as if it were a living entity, separate from the person producing it. Such attribution is not a subjective inference/projection, but an act of meaning-making grounded in basic empathy, linked to the duplicity of movement and corporeality (see 3.1.2). I previously discussed how temporal structures of sound emerge in focal subjectivity, with their own temporality. Still, the temporality of sounds does not imply that they are felt as moving. In focal subjectivity, one can experience outer motion and objects as such in their physical dimension – that is, moving in Merleau-Ponty's "homogeneous" space – through reflection. Nonetheless, the attribution of an-*other* subjectivity requires empathy, thus pertaining to intersubjectivity.

Musical sounds are *perceived as moving*, not as sounds *of* moving objects, and thus not as signs proper, belonging to the following layer of the hierarchy.[12] Moreover, they *express* a sense of movement that the listener shares and "co-shapes" as an "inward experience" (Kim, 2013, p. 165). Clearly, sound is not an actually moving person. Yet, the listener senses movement and feels as if they were moving *with* the sound – in fact, often we find ourselves moving to music instinctually. This is possible because of the felt corporeality of the sound, in relation to the perceived vitality forms. Thus, in addition to – and detaching from, when music is recorded – communication with the musician, we experience a further empathic process with the music.

Specifically, the sense of movement in music listening is grounded in what I earlier termed expressive, inner movement in the living body – linked to Stern's vitality affects – but lacks an outer motion in homogeneous space. Typically, we are used to experiencing the two forms together – as in one's own movement. Yet, inanimate moving objects do not express a sense of living corporeality. Similarly, it is possible to experience inner movement without outer motion.[13]

[11] I say "can" because it is a possible phenomenon, not necessary in all musical experience.

[12] It is irrelevant whether such a way of listening is voluntary: this is merely a description of the experience and its features. The subject can switch their focus to different elements and layers. Identifying the source of sounds could momentarily catch the listener's attention, but it does not prevent them from switching back to experiencing sounds as music in a second moment.

[13] This is also consistent with the relations between the layers in the hierarchy, as empathy can overlap and affect the living body and the structures of subjectivity (temporality of sounds, rhythms, etc.) even without complex structures of intersubjectivity (that is, without all the elements of communication mentioned above, see Section 3.3).

Crucially, this is more than a "metaphorical" description of music *as if* moving. Rather, we perceive music as *moving* and expressive; then, we describe it *as if* someone else was moving with us. In motor intentionality and basic empathy experience is not focal, and the connection between inner movement and the expressivity of vitality forms is experientially motivated, so that one can be expressive of the other. Thus, movement is evoked by the vitality expressed by sounds. Being grounded in this motivated-ness, and devoid of focal aboutness, music is not "about" motion, but it is expressive of movement. Indeed, music is not explicitly about outer motions, nor specific *emotions* (in line with Stern's understanding of these notions, see 3.1.2), at least cross-culturally.

Moving sounds and the virtual person are instances of the kind of experience Merleau-Ponty (1962) described as *aesthetic*: "aesthetic expression confers on what it expresses an existence in itself" (p. 212). As a clarification, I do *not* employ the term "aesthetic" in relation to art, neither in terms of metaphysical/epistemological validity, nor of personal evaluation; rather than having an *a priori* notion of art that would evoke its own modality of experience, I suggest the opposite view. In line with Dissanayake's notion of *artification* (or "making special," see Dissanayake, 2013), communicative processes can be employed to evoke meanings different from the habitual ones. In such cases, features of (inter-)subjectivity are experienced as pertaining to the objects – as also reported in phenomenological investigations of viewing paintings (Starr & Smith, 2021).

Dissanayake's notion resonates with the experience described in this section, perhaps most clearly in Ihde's words: when "aesthetically" listening to music, we detach sounds from their habitual connection to sources (instruments/objects) or communicative content (words in song). Within the new experiential/communicative context of music, sounds are "artified" with a new meaning-making, resulting as expressive of movement.

Finally, let us reconsider once again Husserl's description of melodies as unitary. Investigating the experience of movement in music, Forlè (2016) identifies its core in the notions of ordering and tension – together forming the "teleological relationship" between temporal structures of sounds (pp. 180-181). I suggest that we should separate the two: order emerges in focal intentionality, through recollection of temporal structures, in relation to homogeneous space. Instead, the sense of "teleological" tension/*dynamism* evoked by music is grounded in shared intentionality through the attribution of vitality to sounds – and it is connected to inner movement and "orientated" space. In other words, the perceived vitality/corporeality of sounds and the affective sharing it evokes determine the experienced tension and *unitarity* of musical structures.

It might seem counter-intuitive to locate aesthetic experience in the layer of shared intentionality. Aesthetic experience is undoubtedly grounded in pre-reflective and focal subjectivity, and one might argue that it evokes a specific listening modality, focused on aesthetic goals. Nonetheless, since empathy is critical for experiencing vitality forms from others, and even more so for the "expressed existence" (following Merleau-Ponty) as detached from the sound source and the producing subject, aesthetic experience fits the Semiotic Hierarchy as a sub-layer of shared intentionality.

As mentioned, we can "do" several things with music: I just described the kind of listening that makes one experience sounds as music. Below, I shall consider the meaning-making evoked by music in communication, and how it is related to the listening modality discussed so far.

3.3.3. Communication: signals and systematicity

Communication is the realm of traditional analyses of meaning in terms of "aboutness" and content. Rather than being in contrast with such traditions, my phenomenological approach is parallel to and in dialogue with them. As with Husserl's analysis of time-consciousness, describing the *intentional structure* of communication means to focus on how content manifests experientially – instead of what content is and its specific usages.

Beyond basic levels of intercorporeal interactions and imitation, communication in music concerns the *usage* of sounds with *communicative intent* (see Section 3.3). Usage is essential, as it allows for the sedimentation of *conventions* required for structuring increasingly complex forms of communication. Communicative intent leads to more complex interactions with others, further allowing for inferences, but the communicative possibilities of automatic (stimulus-response) or contextual usage are limited.

The core experience of communicative interactions is *reference*, here defined as the association where an expression [E] stands for an intended object [O] to a given subject [S], be it producer or receiver. To assess the referentiality involved in music, I rely on the definition of *signs* given by Zlatev, Zywiczyński and Wacewicz (2020), based on a qualitative distinction between signs and *signals*, obtained by comparing human and non-human communication. On this definition, signs require the potential for the reflective awareness (in the subject) *of* the signitive process – the reference from [E] to [O] – whereas signals do not. Thus, considering several features related to awareness, volitional control and communicative quality, signs are viewed as establishing relatively stable referential associations between expressions and objects, or *denotations*, as understood by Zlatev et al. (2020). These are detachable from the context and volitionally accessible through memory (see Donald, 1991). In signs, [E] and [O] are identifiable as separate from one another. Signals, instead, rely on implicit "functional" reference: expressions and objects are linked by somewhat stable associations, but their usage is less dependent on volitional control – and, thus, less free to manipulate (Zlatev et al., 2020).

Given this distinction, we need to consider whether cross-cultural musical signification relies on signals or signs. A recent proposal (Perman, 2020) describes musical semiosis as mainly based on purposeful *indices* (indexical signs), assuming a purpose to replicate such signs and, thus, implying a high degree of conventionality within cultures. Nonetheless, it does not consider distinctions between signs and signals, nor between signs and their *grounds* – that is, the *kind* of connection between the expression [E] and its intended object [O] (Sonesson, 2010).

Assuming these distinctions, grounds are common to both signals and signs, and several grounds can be present simultaneously. Thus, the requirement for conventionality within musical cultures leads us to suggest that, with high degrees of convention, music can constitute symbols in Peircean terms – i.e., *signs* whose main ground is conventionality. Yet, Perman's views on the other grounds – iconicity (similarity) and indexicality (causality or contiguity) – do not exclude the possibility for culture-general or less conventionalized musical meanings to depend on signals instead of signs.

I suggest that a further differentiation can be established between signals that are part of a system, and general non-systematic *signalling*: non-human primate calls can constitute systems, with single elements being repeated/repeatable enough to be identified regularly by group members (Sievers & Gruber, 2020; Zlatev et al., 2020), see Figure 2.

General contextual signalling, instead, does not achieve proper reference, because expressions and objects are *not* (fully) differentiated by the subject (be it sender or receiver): such non-systematic associations are mainly in the form of stimulus-response associations. This is also the case with first-time exposure: smoke can only be a sign of fire if I experienced their connection previously. On the other hand, this does not apply to new instances of a recognized and known semiotic system: novel words and neologisms might be correctly interpreted at first exposure, because they arise from within a known system.

In this perspective, associations between expressions and intentional objects in music are not signs, since they do not normally establish context-detachable associations (denotations), and music is hard to use as a tool for deception (Cross, 2009). Music *can* be used to refer to objects/events (Trehub & Trainor, 1992), but these referential associations are mostly context-dependent. Such signals are learnable, distinguishable from one another and systematically repeatable, therefore this kind of musical signification is not a form of pre-systematic signalling. Yet, since these forms of reference are not always evoked in music listening, they do not substantiate the general experience of music.

3.4. Signs: signitive intentionality

In the previous sub-section, I discussed most of the features of signs and their differences from signals. Accordingly, signitive intentionality, emerging from the reflective awareness of the link between [E] and [O], is specifically non-direct: the "contents" of expressions "are not (clearly) given in experience […] because they are primarily intentional acts" (Zlatev, 2018, p. 12). At the same time, signs imply "an irreducible social component" even when they are not used with an explicit desire to communicate, due to their detachability from context – as one's wedding ring can evoke the thought of their spouse.

Within specific cultures, continuous processes of conventionalization can result in higher forms of normativity in music, allowing for the establishment of signs mainly grounded in convention (Zlatev, 2018) – Peircean symbols. This is the case, for instance, with national anthems. As opposed to signals, musical signs can be used more freely and volitionally, and they could be employed to achieve basic forms of deception – e.g., by playing the wrong anthem, or turning it into that of an opposing team. Culture-general music-specific meaning might emerge after (phylogenetically) but on the lower layer, below the sign-threshold. Due to its (diffused) degree of conventionality, music is identified as a semiotic *system*, being perceived as intrinsically intersubjective. Perhaps, despite the lack of stability, musical signals can be employed for reference (see 3.3.3) thanks to the awareness of the sign-function. Still, such signals hold the potential to "turn into" signs within social groups: a musical excerpt could constitute a hymn/anthem within a culture and a foreign listener might be unaware of it; any musical instance could achieve the status of sign within a band's members due to repetition and be employed to refer to a specific object – e.g., a riff from Deep Purple's *Smoke on the Water* to refer to Rudy Ayoub (a YouTube guitarist who purposedly plays that riff in most of his videos). One might hypothesize that musical reference emerges because users are aware of, and used to, reference in general. This implies that we can feel "music!" even if we do not share individual interpretations of it – which depend on interactions between all the layers in the hierarchy and degrees of exposure to a musical culture.

3.5. Symbolic intentionality

The highest layer in the Semiotic Hierarchy regards *symbolic intentionality*, viewed as the possible *relations* in which events/objects in the world can be presented and articulated in our experience. It is symbolic because it manifests through the syntactical articulations of *symbols* – as defined by Zlatev (2018) – in language, but it should not be reduced to a formal (or Chomskyan) notion of grammar. Rather, grammatical constructions in language – "such as predication, modification, conjunction, and subordination" (p. 14) – are acts of meaning-making that allow us to construe and express complex events and relations.

The relational requirements of symbolic intentionality suggest that language might be the only system reaching this layer. The relations involved in musical structures – mainly harmonic connections – seem to be simply formal and, even so, do not correspond to the grammaticality of language (Clarke, 1986). Moreover, since music does not establish *truths* (Packalén, 2008), it is neither possible to lie nor to articulate explicit predications about events.

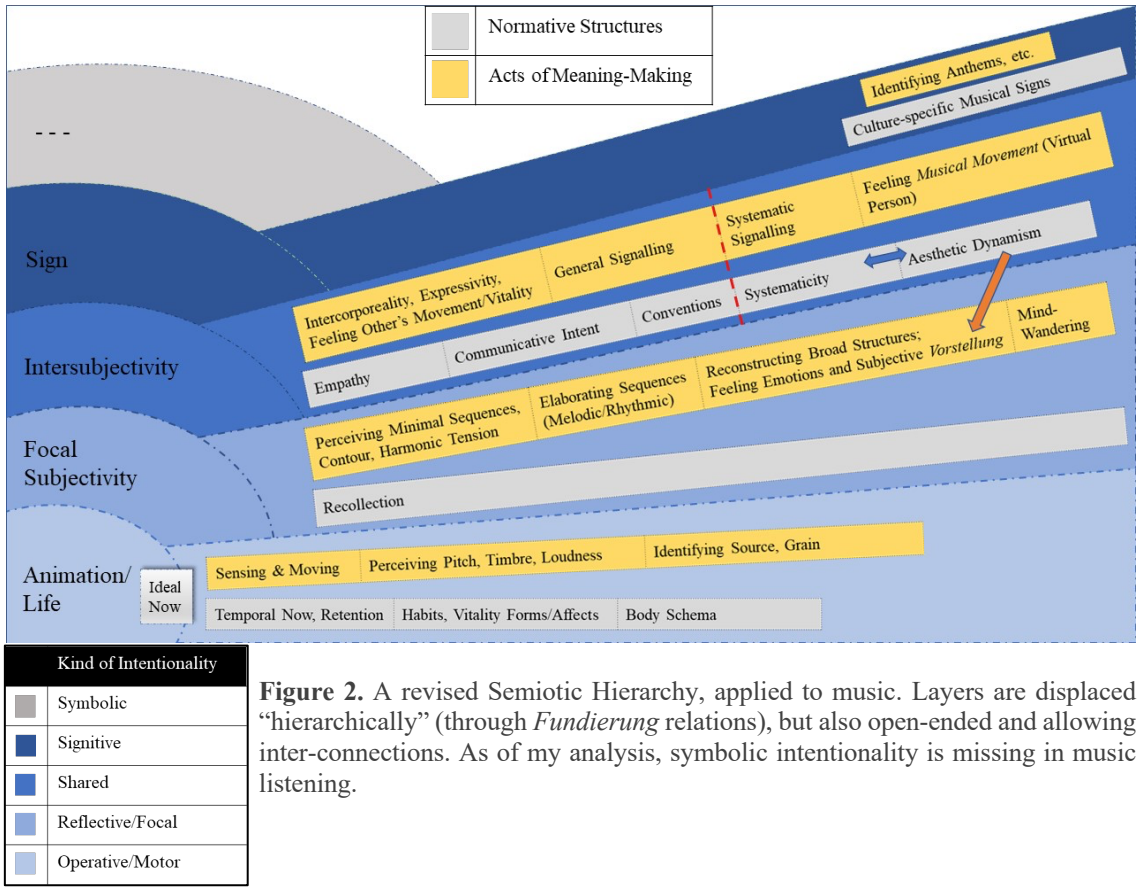

Figure 2. A revised Semiotic Hierarchy, applied to music. Layers are displaced "hierarchically" (through *Fundierung* relations), but also open-ended and allowing inter-connections. As of my analysis, symbolic intentionality is missing in music listening.

4. Interweaving the layers

Figure 2 displays my revised version of the Semiotic Hierarchy model, as applied to music, above all to the aspect of listening to it. As I explained throughout my analysis, layers of experience are not separate; rather, they often interact, resulting in new forms of meaning-making. Let us consider some specific cases of inter-layer connections.

Starting with basic elements, the duality of the living body as "inner" *Leib* and "outer" *Körper* is deeply connected to empathy. As a result, vitality forms derived from one's perception and from empathy are experienced as connected in retentions and protentions, because pre-reflective awareness does not involve attentional focus – therefore, no shift in focus. They are subsequently distinguished focally through reflection and recollection, so that resulting meanings may become focally aware. This allows for the same musical recording to evoke different emotions in different occasions, as the vitality forms evoked by the sounds (and intrinsic expressivity from the musician) are experienced directly in time consciousness as mixed with affects resulting from other sources – such as one's mental status or external conditions. Stern's separation between vitality affects and emotions also implies that specific emotional responses are not central to the experience of music, whereas vitality forms are – that is, I can feel *movement*

in sound even without sharing the musician's emotions, or without feeling any identifiable emotion at all: these occur first on the level of focal subjectivity.

Crucial for the experience of music is the aesthetic experience of sounds as vital and moving (3.3.2). This establishes the ground for the systematicity of musical signaling (3.3.3) – and vice versa: the gradual conventionalization of musical sounds as signals allows for the "artification" process. These two elements sediment hand-in-hand, so that musical sounds are learnt as systematically expressive of inner movement (not specific outer motion). Moreover, in recollection, due to systematicity the listener can *infer* the original source as a living entity (supposedly human) playing the music – or having played, in case of recorded music.[14]

The attribution of vitality to sounds as detached from the musician allows us to experience emotions in relation to music, apart from the emotions of the musician. In poetry and painting I can perceive emotions from the direct facing of the other (in speech or live painting), but I can also feel emotions *not* in relation to that specific person: I do not perceive directly the emotions of the poet the first time I read a poem from someone I never met and I know nothing about. In the same way, I can experience movement and emotions in music apart from those of the musician – although they will blend.

On the other hand, from the offered analysis of empathy in music we could see that musical sounds can express affect – and evoke emotions – even before the subject has "learnt" to experience music as a system. This shows that the connection between musical sounds and emotions does not derive (entirely) from its systematicity: emotions can be communicated across musical cultures.

Finally, affective and emotional responses are potentially related to all layers in the model: sound sequences, interpersonal sharing, communication (through the establishment of the aesthetic dimension) and culture-specific influences. Yet, they are all experienced in a single vitality form: we do not experience retentions over and over to find different elements; rather, we are able to shift focus in recollection to recognize the different elements that compose the one melody/piece we experience.

5. Conclusions

In this article, I adopted a cognitive-semiotic approach to the analysis of meaning in music, due to its potential to elude the limitations of both traditional semiotics and cognitive science. Cognitive semiotics is in favor of trans-disciplinary discussions, allowing for the inclusion of several forms of meaningfulness, such as expanding the notion of meaning to the phenomenological concept of intentionality, thus highlighting the "unity and diversity of meaning" (Zlatev, 2018, p. 14).

More specifically, I applied the Semiotic Hierarchy model (Zlatev 2018; Zlatev & Konderak, 20222) to music listening, to describe the diversity of meaning-making processes involved in music in simultaneity. This tool allowed me to develop a phenomenological analysis of the experience of music listening distributed among several layers of experience, overcoming divisions in other cognitive-semiotic approaches to music – such as perceptual vs. conceptual (see Antović et al., 2020). Thus, I was able to consider the various temporal structures involved in music perception, empathy and interpersonal sharing, and communicative features. Moreover, I highlighted the connections between different layers, showing how this gives rise to further acts of meaning-making.

In the process, I needed to adapt the model in some ways. My version of the Semiotic Hierarchy (Figure 2) includes some updated conceptualizations, clarifying the hierarchical

[14] Clearly, establishing a semiotic system is a cultural fact, happening within a culture. Even considering a "universal" experience of music listening there is a strong influence of culture-specific conventionality, implying that one will get used to experiencing some music in the form of signs – such as specific outer motions according to melodic features (notes "going up/down").

structure as depending on possibilities of interaction rather than on more or less discrete "stages." I also specified my understanding of the interactions between temporality and the living body, allowing for temporality to pervade experience entirely. Stern's notion of *vitality forms* allowed me to describe the motivated connection between movement and affective responses in motor intentionality, while distinguishing between affect (pre-reflective, diffuse) and emotions (reflective, specific). Vitality forms/affects permeate music listening in many ways: from acoustic features, in focal subjectivity, through inter-personal communication with the musician, and in the sense of sharing with music itself.

Furthermore, the connection between motor and shared intentionality led to the addition of the dimension of *aesthetic* experience, which I analyzed as characterizing the specificity of culture-general musical meaning, in relation to the usage of music as a semiotic system. I argued that experiencing music articulates the listener's body in their inner sense of space/time, being grounded in pre-reflective corporeality. Learning to experience music aesthetically makes us feel a sense of movement and vitality in sounds. Accordingly, melodies are experienced through the simultaneous connection of at least three kinds of intentionality, in relation to the temporal flow of awareness and corporeality, recollection of temporal structures and the intersubjective/aesthetic experience of sounds as vital.

Yet, despite constituting a semiotic system, music does not normally evoke *signs* (within the definition I employed from Zlatev et al., 2020), but *signals* – although high degrees of enculturation can lead to musical signs (such as anthems). I also proposed an additional distinction within the layer of communication, separating general signaling (first exposure, entirely context-dependent) from *systematic* signals (repeatable, discrete single elements). Deriving from processes of conventionalization, systematicity gives rise to an implied (sedimented, learnt) sense of sharing and inter-personality.

As a first attempt towards a phenomenological cognitive semiotics of music, this paper only involved first- and second-person methods, obtained through my intuitions and empathetic interpretations of the relevant literature. In the future, these ideas shall be employed for further research, connecting experiential analysis with empirical research while avoiding reductionism, in line with the *phenomenological triangulation* of cognitive semiotics (Pielli & Zlatev, 2020). This triangulation allows for higher degrees of shareability and "extra-subjective" validity, although keeping subjective experience at the core (see also Zlatev, 2015). Since any description can only give some perspective on a phenomenon, investigations based on intuition are a primary step. Then, interpersonal empathetic interactions are required to corroborate individual intuitions, establishing frames for even more detached, possibly quantitative operationalizations.

Accordingly, future research can involve the application of the revised Semiotic Hierarchy to analyze the experience of listening to a musical piece, with the possible addition of qualitative interviews on the same piece. Another line of further investigation might be the comparison with developmental and evolutionary perspectives on cognition and meaning in music (such as Schiavio et al., 2017; Tolbert, 2001). In sum, my adapted Semiotic Hierarchy is by no means exhaustive of all the forms of meaning-making involved in music listening, but it lays the foundation for a transdisciplinary understanding of meaning in music.

Acknowledgments

I am deeply grateful to my friends and colleagues Mingyu Wang, whose suggestions and clarifications both drove me to and helped me understand Husserl's philosophy, and Elio Antonucci, for our stimulating discussions. Many thanks also go to the two anonymous reviewers and the editor-in-chief, Prof. Jordan Zlatev whose comments and suggestions greatly improved the quality of this paper. Finally, my gratitude goes to the Cognitive Semiotics seminar in Lund University for the inspiration and fruitful feedback.

References

Almén, B., & Pearsall, E. (Eds.). (2006). *Approaches to Meaning in Music.* Indiana University Press.

Antović, M., Mitić, J., & Benecasa, N. (2020). Conceptual rather than perceptual: Cross-modal binding of pitch sequencing is based on an underlying schematic structure. *Psychology of Music*, *48*(1), 84–104. https://doi.org/10.1177/0305735618785242

Arbib, M. A. (Ed.). (2013). *Language, Music, and the Brain: A Mysterious Relationship*. The MIT Press. https://doi.org/10.7551/mitpress/9780262018104.001.0001

Burling, R. (2005). *The Talking Ape*. Oxford University Press.

Clarke, D. (2011). Music, phenomenology, time consciousness: meditations after Husserl. In D. Clarke, & E. F. Clarke (Eds.), *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 1-28). Oxford University Press. https://doi.org/fx9gdj

Clarke, E. (1986). Theory, Analysis and the Psychology of Music: A Critical Evaluation of Lerdahl, F. and Jackendoff, R., A Generative Theory of Tonal Music. *Psychology of Music*, *14*(1), 3–16. https://doi.org/10.1177/0305735686141001

Clarke, E. (2011). Music perception and musical consciousness. D. Clarke, & E. F. Clarke (Eds.), *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 193-214). Oxford University Press. https://doi.org/fx5nt5

Clarke, E. (2017). Empathic entanglements: Music, motion, dance. In P. Veroli, & G. Vinay (Eds.), *Music-Dance: Sound and Motion in Contemporary Discourse* (pp. 191-206). Routledge.

Clarke, E. (2019). Empathy and the ecology of musical consciousness. In R. Herbert, D. Clarke, & E. Clarke (Eds.), *Music and Consciousness 2: Worlds, Practices, Modalities* (pp. 71-92). Oxford Academic. https://doi.org/10.1093/oso/9780198804352.003.0005

Cochrane, T. (2010). Using the Persona to Express Complex Emotions in Music. *Music Analysis*, *29*(1/3), 264–275. http://www.jstor.org/stable/41289728

Cross, I. (2009). The Evolutionary Nature of Musical Meaning. *Musicae Scientiae*, *13*(2_suppl.), 179-200. https://doi.org/10.1177/1029864909013002091

Cross, I. (2011). Music as a social and cognitive process. In P. Rebuschat, M. Rohrmeier, J. Hawkins, & I. Cross (Eds.), *Language and Music as Cognitive Systems*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199553426.003.0033

Cross, I., & Tolbert, E. (2016). Music and meaning. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford Handbook of Music Psychology* (2nd ed., pp. 33-46). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198722946.013.7

Cross, I., & Tolbert, E. (2020). Epistemologies. In T. McAuley, N. Nielsen, J. Levinson, & A. Phillips-Hutton (Eds.), *The Oxford Handbook of Western Music and Philosophy* (pp. 265-282). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199367313.001.0001

Deutsch, D., Hentohorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *The Journal of the Acoustical Society of America*, *129*, 2245-2252. https://doi.org/10.1121/1.3562174

Devylder, S., & Zlatev, J. (2020). Cutting and Breaking Metaphors of the Self and the Motivation and Sedimentation Model. In A. Baicchi (Ed.), *Figurative Meaning Construction in Thought and Language* (pp. 253-282). Benjamins. https://doi.org/10.1075/ftl.9.11dev

Dissanayake, E. (2013). Genesis and development of "Making Special": Is the concept relevant to aesthetic philosophy? *Rivista di Estetica*, *54*, 83-98. https://doi.org/10.4000/estetica.1437

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of human culture.* Harvard University Press.

Forlè, F. (2016). Movement in Music. An Enactive Account of the Dynamic Qualities of Music. *HUMANA.MENTE Journal of Philosophical Studies*, *9*(31), 169-185. https://www.humanamente.eu/index.php/HM/article/view/50

Gabrielsson, A. (2011). *Strong Experiences with Music: Music is Much More Than Just Music*. Oxford University Press. https://doi.org/fxzst4

Gallagher, S. (1998). *The Inordinance of Time*. Northwestern University Press.

Godøy, R. I. (2011). Sound-action awareness in music. In D. Clarke, & E. F. Clarke (Eds.), *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 231-244). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199553792.003.0069

Guareschi, C. (2019). *Merleau-Ponty's philosophy of nature and intercorporeality: an embodied model for contemporary environmental aesthetics* [Doctoral dissertation, University College Cork]. http://creativecommons.org/licenses/by-nc-nd/3.0/

Hilpinen, R. (2015). Conception, sense, and reference in Peircean semiotics. *Synthese*, *192*, 991-1018. https://doi.org/10.1007/s11229-013-0326-9

Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal Time (1893-1917)* (J. B. Brough, Trans.). Kluwer Academic Publishers.

Ihde, D. (2007). Listening and Voice. Phenomenologies of Sound (2nd ed.). Suny Press.

Kim, J. H. (2013). Shaping and Co-Shaping Forms of Vitality in Music: Beyond Cognitivist and Emotivist Approaches to Musical Expressiveness. *Empirical Musicology Review*, *8*(3-4), 162-173. https://doi.org/10.18061/emr.v8i3-4

Koelsch, S. (2011). Toward a Neural Basis of Music Perception – A Review and Updated Model. *Frontiers in Psychology*, *110*. https://doi.org/10.3389/fpsyg.2011.00110

Koelsch, S. (2019). Music and the brain. In P. J. Rentfrow, & D. J. Levitin (Eds.), *Foundations in Music Psychology: Theory and Research* (pp. 407-458). The MIT Press.

Krueger, J. (2019). Music as Affective Scaffolding. In R. Herbert, D. Clarke, & E. Clarke (Eds.), *Music and Consciousness 2: Worlds, Practices, Modalities* (pp. 55-70). Oxford University Press. https://doi.org/10.1093/oso/9780198804352.003.0004

Kühl, O. (2008). *Musical semantics*. Peter Lang. https://doi.org/10.3726/978-3-0351-0271-0

Larrabee, M. J. (1989). Time and Spatial Models: Temporality in Husserl. *Philosophy and Phenomenological Research*, *49*, 373-392. https://doi.org/10.2307/2107794

Leman, M. (2007). *Embodied music cognition and mediation technology*. The MIT Press. https://doi.org/10.7551/mitpress/7476.001.0001

Levinson, J. (2006). Musical Expressiveness as Hearability-as-Expression. In M. Kieran (Ed.), *Contemporary Debates in Aesthetics and the Philosophy of Art* (pp. 192-206). Blackwell Publishing. https://doi.org/10.1093/acprof:oso/9780199206179.003.0007

Livingstone, R. S., & Thompson, W. F. (2009). The emergence of music from the theory of mind. *Musicae Scientiae*, *13*(2 suppl), 83–115. https://doi.org/10.1177/1029864909013002061

Malloch, S., & Trevarthen, C. (2018). The Human Nature of Music. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01680

Mazzola, G. (2002). *The topos of music*. Birkhauser Verlag. https://doi.org/10.1007/978-3-0348-8141-8

Mazzola, G. (2003). Semiotic Aspects of Musicology: Semiotics of Music. In R. Posner, K. Robering, & T. A. Sebeok (Eds.), Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur / Handbook on the Sign-Theoretic Foundations of Nature and Culture. 3. Teilband (pp. 3119–3188). De Gruyter Mouton. https://doi.org/10.1515/9783110194159

Mehr, S. A., Singh, M., Knox D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and Diversity in Human Song. *Science*, *366*, Article eaax0868. https://doi.org/ggdvjp

Merleau-Ponty, M. (1962). *Phenomenology of Perception* (C. Smith, Trans.). Routledge & Kegan Paul. (Original work published 1945) https://doi.org/10.4324/9780203994610

Merleau-Ponty, M. (1968). *The Visible and the Invisible* (A. Lingis, Trans.). Northwestern University Press. (Original work published 1964)

Mirigliano, R. (2011). The sign and music: A reflection on the theoretical bases of musical semiotics. In E. Tarasti (Ed.), *Musical Signification: Essays in the Semiotic Theory and Analysis of Music* (pp. 43-62). De Gruyter Mouton. https://doi.org/10.1515/9783110885187.43

Miu, A., & Vuoskoski, J. (2017). The social side of music-listening: Empathy and contagion in music-induced emotions. In E. King, & C. Waddington (Eds.). *Music and Empathy* (pp. 124-138). Routledge. https://doi.org/10.4324/9781315596587-6

Monelle, R. (1991). *Linguistics and semiotics in music*. Routledge. https://doi.org/10.4324/9781315076942

Montague, E. (2011). Phenomenology and the 'hard problem' of consciousness and music. In D. Clarke, & E. F. Clarke (Eds.), *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 29-46). Oxford University Press. https://doi.org/fz35vf

Packalén, E. (2008). Music, Emotions, and Truth. *Philosophy of Music Education Review*, *16*(1), 41-59. https://doi.org/10.2979/pme.2008.16.1.41

Perman, T. (2020). Musical meaning and indexicality in the analysis of ceremonial mbira music. *Semiotica* (236-237), 55-83. https://doi.org/10.1515/sem-2018-0057

Pielli, L., & Zlatev, J. (2020). The cyborg body: Potentials and limits of a body with prosthetic limbs. *Cognitive Semiotics*, *13*(2). https://doi.org/10.1515/cogsem-2020-2033

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

Schiavio, A., van der Schyff, D., Kruse-Weber, S., & Timmers, R. (2017). When the Sound Becomes the Goal. 4E Cognition and Teleomusicality in Early Infancy. *Frontiers in Psychology*, *8*, 1585. https://doi.org/10.3389/fpsyg.2017.01585

Seifert, U. (2020). Relating the evolution of Music-Readiness and Language-Readiness within the context of comparative neuroprimatology. In M. A. Arbib (Ed.), *How the Brain Got Language – Towards a New Road Map* (pp. 86-101). John Benjamins Publishing. https://psycnet.apa.org/doi/10.1075/is.17035.sei

Sheets-Johnstone, M. (2011). *The Primacy of Movement* (2nd ed.). John Benjamins Publishing. https://psycnet.apa.org/doi/10.1075/aicr.82

Sheets-Johnstone, M. (2012). Fundamental and inherently interrelated aspects of animation. In A. Foolen, U. Lüdtke, T. Racine, & J. Zlatev (Eds.), *Moving Ourselves, Moving Others: Motion and emotion in intersubjectivity, consciousness and language* (pp. 29-56). John Benjamins Publishing. https://doi.org/10.1075/ceb.6.02she

Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proc. Natl. Acad. Sci. U.S.A.*, *2, 110*(1), 70-75. https://doi.org/10.1073/pnas.1209023110

Sievers, C., & Gruber, T. (2020). Can Nonhuman Primate Signals be Arbitrarily Meaningful like Human Words? An Affective Approach. *Animal Behavior and Cognition*, *7*(2). https://doi.org/10.26451/abc.07.02.08.2020

Sokolowski, R. (2008). Husserl's Discovery of Philosophical Discourse. *Husserl Studies*, *24*, 167-175. https://doi.org/10.1007/s10743-008-9043-5

Sonesson, G. (2010). From Mimicry to mime by Way of Mimesis: Reflections on a General Theory of Iconicity. *Sign Systems Studies*, *38*, 18–66. https://doi.org/10.12697/SSS.2010.38.1-4.02

Sonesson, G. (2012a). The foundation of cognitive semiotics in the phenomenology of signs and meanings. *Intellectica*, *58*(2), 207-239. https://doi.org/10.3406/intel.2012.1105

Sonesson, G. (2012b). The meanings of structuralism: Considerations on structures and Gestalten, with particular attention to the masks of Levi-Strauss. *Segni e comprensione*, *XXVI*(78), 84-101. https://doi.org/10.1285/I18285368AXXVIN78P84

Starr, R. A., & Smith, J. A. (2021). "People are gazing" — An Interpretative Phenomenological Analysis of Viewing Velázquez. *Art & Perception*, *9*(3), 241-259. https://doi.org/10.1163/22134913-bja10027

Stern, D. (1985). T*he Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. Basic Books. https://doi.org/10.4324/9780429482137

Stern, D. (2010). Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy and Development. Oxford University Press. https://doi.org/jjkz
Teie, D. (2016). A Comparative Analysis of the Universal Elements of Music and the Fetal Environment. *Frontiers in Psychology*, *7*, Article 1158. https://doi.org/ghk534
Thompson, E. (2001) Empathy and consciousness. *Journal of Consciousness Studies*, *8*(5–7), 1–32.
Thompson, E. (2007). Mind in life: Biology, phenomenology, and the sciences of mind. Belknap Press/Harvard University Press.
Tolbert, E. (2001). Music and Meaning: An Evolutionary Story. *Psychology of Music*, *29*, 84-94. https://doi.org/10.1177/0305735601291006
Trehub, S. E., & Trainor, L. J. (1992). The Development of Referential Meaning in Music. *Music Psychology*, *6*(4), 455-470. https://psycnet.apa.org/doi/10.2307/40285565
van der Schyff, D., & Krueger, J. (2019). Musical Empathy, From Simulation to 4E Interaction. In A. F. Correa (Ed.), *Music, Speech & Mind* (pp. 73-107). ABCM.
Watt, R. J., & Ash, R. L. (1998). A psychological investigation of meaning in music. *Musicae Scientiae*, *2*(1), 33–53. http://dx.doi.org/10.1177/102986499800200103
Zahavi, D. (2001). Beyond Empathy. Phenomenological approaches to intersubjectivity. *Journal of Consciousness Studies*, *8*, 151-167.
Zahavi, D. (2003). Inner time-consciousness and pre-reflective self-awareness. In D. Welton (Ed.), *The New Husserl: A critical reader*, (pp. 157–180). Bloomington: Indiana University Press.
Zlatev, J. (2002). Meaning = Life (+ Culture): An outline of a unified biocultural theory of meaning. *Evolution of Communication*, *4*(2), 255–299. https://doi.org/dvfwnn
Zlatev, J. (2009). The semiotic hierarchy: Life, consciousness, signs and language. *Cognitive Semiotics*, *4*(supplement), 169–200. https://doi.org/jjk7
Zlatev, J. (2012). Prologue: Bodily motion, emotion and mind science. In A. Foolen, U. Lüdtke, T. Racine, & J. Zlatev (Eds.), *Moving Ourselves, Moving Others* (pp. 1-26). John Benjamins Publishing. https://doi.org/10.1075/ceb.6.01zla
Zlatev, J. (2015). Cognitive semiotics. In P. Trifonas (Ed.), *International handbook of semiotics* (pp. 1043–1067). Springer. https://doi.org/10.1007/978-94-017-9404-6
Zlatev, J. (2018). Meaning making from life to language: The semiotic hierarchy and phenomenology. *Cognitive Semiotics, 11*(1). https://doi.org/10.1515/cogsem-2018-0001
Zlatev, J., & Konderak, P. (2022). Consciousness and semiosis. In J. Pelkey (Ed.), *Bloomsbury Semiotics: History and Semiosis* (pp. 169-192). Bloomsbury Publishing Plc. http://dx.doi.org/10.5040/9781350139312.ch-8
Zlatev, J., Zywiczyński, P., & Wacewicz, S. (2020). Pantomime as the original human-specific communicative system. *Journal of Language Evolution*, *5*(2), 156–174. https://doi.org/10.1093/jole/lzaa006

Author address
gabriele.giacosa@gmail.com

About the author
Gabriele Giacosa is a PhD candidate at the University of Cologne, and a collegiate member of the a.r.t.e.s. Graduate School. He is working on a transdisciplinary project at the intersection of systematic musicology, phenomenology, and cognitive semiotics, under the supervision of Prof. Uwe Seifert (University of Cologne), Prof. Jordan Zlatev (Lund University), and Prof. Shigeru Taguchi (Hokkaido University). Gabriele holds a R.MA in Musicology from Utrecht University and a BA in DAMS (Music) from the University of Turin. His research interests span from the philosophy of music and aesthetics to cognitive science, with a focus on music, language and meaning.