



# Listeners' Bodies in Music Analysis: Gestures, Motor Intentionality, and Models\*

Mariusz Kozak

NOTE: The examples for the (text-only) PDF version of this item are available online at:

<http://www.mtosmt.org/issues/mto.15.21.3/mto.15.21.3.kozak.php>

KEYWORDS: motion-capture, gesture, analysis, Merleau-Ponty, motor intentionality, Neuwirth, Carter, Adès

ABSTRACT: In this article I demonstrate how listeners understand musical processes with their bodies, and how their gestures can be used to build analytical models. Specifically, I draw on the phenomenology of Maurice Merleau-Ponty to argue that situated, active listeners project their motor intentional gestures inside music, where they reconstitute the very nature of musical space and its objects according to their own unique perspective. Rather than passively reflecting gestures of performers, these listeners use their own bodily states to create the structure and meaning of music. I illustrate how those states can be mobilized for analysis by taking quantifiable features of gestures—acceleration and temporal profiles—as models of musical structure, and by using those models as a basis for analytical narratives. I focus on three pieces—Olga Neuwirth's *Vampyrotheone*, Elliott Carter's *ASKO Concerto*, and Thomas Adès's *Living Toys*—in which motion-capture studies revealed the different roles of listeners' gestures in organizing musical experience.

*Received February 2015*

---

## I. Introduction

[1.1] Mahler's symphonies are rarely, if ever, discussed in the context of embodiment of any kind. Suspended in the space between program and absolute music, their meaning seems to emerge from a structure that invites a kind of transcendental hearing, one that aims to supersede earthly, everyday experience and transpose the listener into some other realm.<sup>(1)</sup> Yet there is at least one Mahlerian moment in which I am excruciatingly aware of my body in relation to the music: the famous "Great Summons" (*Grosse Appel*) in the Finale of his Second Symphony ("Resurrection"). If the passage is performed right, I feel an almost unbearable tension as the final quivers of the nightingale's death song intermingle with the heavenly post-apocalyptic trumpets, dissolving into a thick, pregnant silence just before the long-awaited entrance of the chorus. The tautness of this moment—remarkably acute when the singers remain seated for the first few lines of Klopstock's text, a posture that conceals their performative intent and magnifies the suspense—is mirrored in the physical strain of my body, as I find myself leaning forward expectantly, at the same time catching the last echoes of the Summons and anticipating the breathtaking Apotheosis that will take me to the end of the piece. Muscles, tendons, and organs strain in stillness, fighting to keep inert. This is an agitated repose, and its pressure can only be released in the brilliant roar of the ensemble in the final measures of the symphony some ten minutes later.

[1.2] In this article I explore some ways in which an experience such as this can be harnessed for the purposes of musical analysis. More specifically, I am interested in conceptually grounding the role of listeners' mobile, active bodies in their engagement with music, and inquiring into how those bodies can be said to constitute a fundamental level of musical understanding. To be sure, the musical encounter I described above is hardly unique, and treads down well-worn paths by foregrounding the function of an embodied, situated participant in the formation of musical meanings. Indeed, in some ways it provides an entry point for what Fisher and Lochhead (2002) call an "analysis from the body." Growing significantly in

number over the past decade or so, entreaties for precisely this kind of engagement with music have come from different directions, unified by a common thesis that music must be understood as something we *do*, and not something that is given, or revealed, to us.<sup>(2)</sup> In music theory, the call has been answered with an expanding corpus of scholarship that shifts the attention of analysis from the finished musical “object” to the emerging process of performance. Either by presenting first-person phenomenological accounts or extrapolating some basic movement features of the instrumentalist (typically a pianist), such studies of musical embodiment try to determine how the performing body participates in the emergence of musical meanings that are captured neither by sounds alone, nor by analytical tools designed to deal exclusively with those sounds.

[1.3] Highly influential in shaping the discourse in that regard was Suzanne Cusick’s (1994) vision of an embodied music theory that acknowledged how bodies mediate analytical descriptions of music. Although specifically exploring the role of gender in music production, Cusick proposed that theorists consider different ways in which meanings issue from the very fact that music necessitates performance, and so delve critically into hearing the body that is enacted *in* it. Following Cusick’s lead, Fisher and Lochhead (2002) used their own experiences as instrumentalists to examine the possible analytical work of gestures in practicing, rehearsing, and performing Joan Tower’s *Fantasy (those harbor lights)* and the third movement of Johannes Brahms’s Sonata for Clarinet and Piano in E-flat major, op. 120, no. 2. On their account, the privileging of the performers’ bodies is justified for two reasons: it is precisely those bodies that visibly move when playing, and it is these very movements that suggest “an underlying experiential correspondence” with the qualitative and affective aspects of musical sounds (2002, 47). From a less subjective perspective, Mead (2002) explores the structural role of hand-crossings in the second movement of Anton Webern’s Variations for Piano op. 27. Here, the unusual physicality of performance is written into the score—or rather, its unusualness is made conspicuous—in such a way that a structural layer emerges quite separately from that suggested by pitch and rhythmic materials. Namely, while the latter are governed by a two-voice canon, the hands embody articulations of motivic returns over and above musical form.

[1.4] More recent examples demonstrate increasingly varied methods of incorporating corporeal experiences of performers into music analysis by combining them with existing analytical technologies. For example, Koozin (2011) draws on transformational analysis to map the motion of harmonic patterns through a guitar fretboard, thus illustrating how musical materials in pop-rock guitar music are intimately linked to their embodied instantiations and the socio-cultural meanings from which they emerge. In a phenomenological account, Montague (2012) examines how a single performance gesture—an extension of the fifth finger, and the resultant spreading out of the whole hand—in Chopin’s Étude in A-flat major, op. 25, no. 1, is tied to the expressive content of the piece. Basing the analysis on his own experience, he posits that this gesture organizes melodic and harmonic materials in a way that creates a narrative meaning that emerges from the performing body, and not just from relationships between sounds. Finally, taking a more quantitative approach to movement analysis, MacRitchie and colleagues (2013) illustrate that pianists employ expressive gestures to convey certain structurally significant musical elements, such as phrases and other kinds of grouping, or melodic and harmonic climaxes. This suggests that performers physically embody musical structure that they communicate to listeners, leading the authors to posit that “overall motion in a performance arises from the performer’s own representation of the musical structure” (2013, 103).<sup>(3)</sup>

[1.5] The selective overview presented above can hardly do justice to the many ways in which these and other studies have illuminated fascinating relationships between bodies and music. However, because they draw exclusively on gestures of performers, these approaches simultaneously imply a certain level of passivity on the part of the embodied listener.<sup>(4)</sup> This is to say that they can only account for what the *performer* experiences; they are unable to address directly the perspective of *listeners* beyond conjecture about how performers’ bodily states (gestures, postures) translate to, or are picked up by, the audience. As such, my opening paragraph would be unintelligible in this context, because nothing in my description of anticipatory rigor in Mahler invokes gestures that come to bear directly on musicians’ production of sound.<sup>(5)</sup> While Fisher and Lochhead write that “hearing entails a bodily enactment of musical meaning that links listeners, performers, and creators in the same musical enterprise” (2002, 46), they neither explore how these perspectives are linked, nor extend this link to an active bodily engagement with musical performances that are unseen. Indeed, they explicitly claim that when deprived of the visual component of performance, listeners’ “performative enactment of musical meaning relies on a prior backdrop of experience that allows [them] to imaginatively engage the physical activities that went into its production” (2002, 47). Similarly, Brown suggests that sounds themselves are “scattered relics of past, embodied actions” (2006, 43), a *memento mori* of embodiment that listeners experience by moving through the world of physical objects.<sup>(6)</sup> Music from this perspective offers a strikingly visual element in which we “see ourselves sounding” in the way in which music “makes prolongation of our own sounding bodies.” In other words, the listener’s body re-enacts the memory of original sound production; it picks up the relics and visually projects itself into music as if possessed by the spirit of the performer. Or, as Godoy suggests, “massive ecological knowledge of sound production means that listeners . . . have a repertoire of sound-producing gestures so that in situations where there are no visible musicians . . . , the listeners may mentally recreate the choreography of sound-producing gestures” (2010, 106).<sup>(7)</sup>

[1.6] Altogether, claims such as these might lead to an observation that, for listeners, music is always embodied only contingently (for example, when dancing, or when viewing a performance, as Fisher and Lochhead suggest), and that performers’ gestures are far more relevant as objects of analytical interest. Yet we should be cautious in reaching this conclusion. For one thing, listeners come from diverse backgrounds and might know nothing about how sounds are

produced on a particular instrument (by what bodily and technical means) and what gestural constraints and opportunities this instrument affords.<sup>(8)</sup> Moreover, instead of actively participating in making sense of what they hear, listeners' bodies become inert historical backdrops for more traditionally conceived cognitive capacities of imagination and conceptualization: at some point in the past (usually in infancy) they served to connect the physicality of sound production with the perception of its effects, but this role is not actively deployed at the present moment of experiencing music performed by someone else.<sup>(9)</sup> This means that these bodies are quite unnecessary—or at least extraneous—to the formation of musical meanings, in which case we might describe the listening situation as quasi-embodied at best. Finally, the purported transference between performers and listeners impedes a more profound inquiry into the ways in which listeners' bodily experiences meaningfully structure their understanding of musical sounds in a manner that can be useful to analysis, because those very experiences are always auxiliary to the gestures of performers, coupled in a seemingly causal fashion to the various methods of sound production. Thus, the body of the listener turns into either a quiescent mirror of someone else's actions or an automaton that impulsively reacts to external stimuli.

[1.7] Still, there are circumstances that make it difficult to incorporate listeners' bodies into music analysis, even if we acknowledge their autonomous function in the formation of musical meanings. Namely, gestures of listeners are either attenuated, as in the case of most engagements with Western art music, or too fleeting and idiosyncratic to contribute to a rigorous theoretical inquiry. By contrast, performers' actions are both visible, and generally obtain from one instance to another. For example, Montague's finger extension or Koozin's fretboard patterns can form a very specific vocabulary that is more or less enduringly anchored to definite sequences of sounds, a vocabulary that is constrained by the physical properties of instruments to which these gestures are directed and by the biomechanics of players' bodies. Consequently, these gestures exhibit a kind of stability that not only allows external observers to study them through an objective lens, but also lets analysts easily share their findings with the broader community of interested parties.

[1.8] How, then, are we to make sense of listeners' bodies as contributing to musical understanding, doing so in a way that is secured just enough to allow for at least a modicum of intersubjectivity? How can we find an appropriate balance between what Carolyn Abbate (2004) calls “drastic” and “gnostic” engagements with musical materials, a balance that retains in full view both the vulnerability of the body and the fixity of sound objects? Taking seriously the argument that music is something that even listeners *do*, and something they do fully embodied and without the contingency of performance, I propose using their moving bodies as the foundation for a kind of musical understanding that *could* productively underlie a rigorous music analysis by complementing, instead of merely reflecting, analysis grounded in performers' gestures. The goal of my inquiry is to chart a path toward an analysis that uses listeners' bodies as one of the tools in its arsenal, which I will demonstrate with analytical snapshots where motion-capture technology was used to suggest musical moments that may be of structural interest. In forging through the theoretical and empirical landscape, three questions will lurk in the background: (1) How do listeners' bodies do analytical work? (2) In what meaningful ways do those bodies alter musical space and the constitution of its objects? (3) How can we as analysts use those bodies in ways that are sufficiently objective and rigorous to effect our ability to share our findings with others, but sensitive enough to individual idiosyncrasies not to stifle the contingency of musical meaning?

## II. Musical Insiders and “Characteristic Gestures”

[2.1] How *do* listeners' bodies participate in generating viable analyses of music? To begin addressing this question, I propose drawing on an existing music theoretical model in which bodies are already implicated in manifesting relationships between musical objects: David Lewin's theory of transformations. This theory—which I take to span not only his monographs (Lewin 1987, 1993), but also his widely celebrated essay on music perception (Lewin [1986] 2006)—is arguably the most clearly articulated position with regard to the phenomenology of music analysis, and as such bears directly on experience.<sup>(10)</sup> The purpose of my gambit is to extend the notions of perception in Lewin's approach in a way that ends up encompassing the active constitution of musical objects by an embodied listener. To do that, in this section I formulate the problem of “bodily analysis” as a concern with musical insideness, or with placing the listener in a position that generates phenomena of analytical interest “from within” (so to speak) the musical context.

[2.2] In his highly influential *Generalized Musical Intervals and Transformations* (1987), Lewin claims that music analysts typically consider themselves as observers looking at an external “musical space.”<sup>(11)</sup> To elaborate this idea further, we can note that such a space supplies a grid of well-defined points that are occupied by musical “elements.” These elements, in turn, can be objectively measured using intervals, construed as distances between unique points in the space. For example, the pitches C5 and G5 occupy specific locations in “pitch space,” and one way in which the distance between them can be expressed is as an interval of a perfect fifth. This interval obtains regardless of the specific sequence in which those pitches are presented. Importantly, pitch space thus conceived is absolute, and so an interval between C5 and G5 is exactly the same as an interval between F3 and C4, or between A6 and D6. In other words, the phenomenal qualities—or *qualia*—of the pitches involved in the measurement are irrelevant to the measurement itself.<sup>(12)</sup> Or, to put it more positively, all elements of pitch space are phenomenally equivalent with regard to the process of calculating distances between them. This suggests that the space itself is uniform and exists separately from its contents, which subsequently become “idealized abstractions” on a Cartesian plane. Not only that, the space is thoroughly generic, as different sets of elements can take up position on the same grid. According to Lewin, whenever musical properties—not just pitches and rhythms, but also harmonies, time points, and even

timbres—can be modeled using mathematical groups, the relationship between them can be formally characterized using intervals.<sup>(13)</sup>

[2.3] Lewin contrasts the position of a detached observer with that of a musical insider, whose attention is focused on holistic “contextually articulated phenomena” rather than elemental sonic particles. He writes: “Given locations *s* and *t* in our space [N.B. the space could be comprised of any kind of elements, not limited to pitches—auth.], this attitude does not ask for some observed measure of extension between reified ‘points’; rather, it asks: ‘If I am *at s* and wish to get to *t*, what characteristic gesture . . . should I perform in order to arrive there?’” (Lewin 1987, 159). Whereas an outsider operates wielding a ruler notched with intervals, the tools of the insider are transformations. Elsewhere he adds that a transformation is “something one *does* to a *Klang*, to obtain another *Klang*” (1987, 177). It is this performative aspect of transformations that concerns us here: placing listeners *inside* music, alongside musicians and dancers. One consequence of this placement is that it entails reconfiguring not just musical objects themselves (from atomistic sounds to holistic *Klang*s), but also the space in which they are conceptualized (from a Cartesian plane to something closer to experience).

[2.4] Perhaps the most obvious way to interpret Lewin’s claims of musical insideness, of doing things, and of “characteristic gestures” is metaphorically. After all, sounds are not physical entities that one can manipulate in a way that would causally affect another sound; there is nothing in sound that one can do anything to in the same way that one can grasp a book, cradle the hand of a loved one, or mold a lump of clay into a bust of Beethoven. Nor is there anything that a sound can literally do on its own, without the aid of metaphors.<sup>(14)</sup> Steven Rings addresses this very issue in his monograph *Tonality and Transformation* and suggests that the concept of *intentionality*—understood in Husserlian terms as a property of consciousness as it is directed toward an object or event, i.e. as it is “about” something—may help clarify the proposal that “transformations model first-person actions of some kind” (2011, 104). As he sees it, transformations model a process of hearing by specifying some kind of a relationship that one sound bears to another. On this view, a transformation is an intentional attitude toward a presently sounding event that the listener needs to adopt in order to conceptualize this same event as sounding in a particular relationship to what comes after (in networks that model temporal orientation, or Rings’s “event networks”), or to what governs a number of non-sequential events (in networks that model atemporal relationships, or Rings’s “spatial networks”).

[2.5] For Rings, Lewin’s statement concerns “doing something” metaphorically: we do not perform real actions on the sounds we hear. And indeed, Lewin keeps deferring the perspective of the listener, claiming instead that the transformational attitude is that of an “idealized dancer and/or singer,” while leaving the listener/analyst to look on from the outside (1987, 159). In other words, to “do something” to music, the listener needs to become *like* a dancer or a singer, with the implication that listeners do not, in fact, do anything to music just “as is”—something that I will contest below. However, as Brian Kane suggests, this attitude reveals a problem in Lewin’s thought concerning the relationship between Husserlian and post-Husserlian phenomenology. According to Kane, “there is evidence to suggest that Lewin’s commitments point in a direction away from Husserlian intentionality, toward post-Husserlian embodiment,” but it is not a direction that Lewin himself pursued (2011, 34). Quite the opposite, by regarding perception as something that simply happens to an inert, disengaged subject, Lewin explicitly argued that it is inadequate as a foundation for a theory of music because it closes off productive avenues of interpretation.<sup>(15)</sup> Pointing instead to the creative power of musical behavior constituted in performance and composition (among others<sup>(16)</sup>), he cleared the way for the possibility of linking musical understanding with the generative multiplicity of what he whimsically dubbed “noodling” and “fooling around” (Lewin [1986] 2006, 97).

[2.6] In light of this tension between a perceiving organism and a conceiving musician, Kane actually sees two options for following a Lewinian connection between phenomenology and a theory of music. Option one is to reconcile Lewin’s claims of embodiment with his formalized analytical technologies, which is the tactic that Rings uses to abandon the notion of literal embodiment. Option two is to bracket the formalism and go ahead and flesh out Lewin’s critique of Husserlian phenomenology in post-Husserlian terms, something that could be accomplished by looking at embodiment through a lens of “second-wave” phenomenologists like Maurice Merleau-Ponty, whose *Phenomenology of Perception*—originally published in French in 1945 as *Phénoménologie de la perception*—fundamentally altered the landscape of contemporary studies on the relationship between a perceiving subject and its environment.<sup>(17)</sup> As I will explore in more detail in Part IV of this article, one thing that falls out of this approach is a reconceiving of perception as a process that is integrally yoked with action, and one that, in this case, actively constitutes the listener’s experience of musical objects.

[2.7] In the following two sections of this article, I flesh out a phenomenological critique in terms of Merleau-Ponty’s theory of perception, which entails taking a favorable stance on Lewin’s views and extending them through the notion of listeners’ bodily engagement with musical sounds. There, we will see that the literal source of Lewin’s characteristic gestures is found in the material body. In other words, such gestures are not abstract shapings of some metaphysical musical energy—a kind of symbolic action that an analyst might effect in lieu of *actually* performing the music—but real movements of real, embodied listeners.<sup>(18)</sup> They are the non-figurative, tangible source of *doing* something to a *Klang*, one that can help us mobilize the listening body as an analytical tool.

[2.8] However, it is not enough to simply assert that such a material body plays a non-metaphorical role in the constitution of musical spaces and objects inhabiting those spaces. If listeners are to transform into figures on par with an “idealized dancer

and/or singer,” we must show that they are *de facto* capable of performing gestures in order to arrive at certain points in the music. Grounding the argument in Merleau-Ponty’s concepts of *motor intentionality* and *spatiality of situation* provides us with the conceptual mechanism to do just that, whereby listeners’ gestures can be said to assimilate the structure of music into those listeners’ “substance,” such that this very structure directly regulates the subjects’ movements (Merleau-Ponty [1945] 2012, 134).<sup>(19)</sup> In what the philosopher calls a dialogue between the subject and the object, the listener draws together the meanings diffused throughout music, while music does the same with the listener’s bodily comportment. Both the active listener and the music reciprocally direct one another, the former constituting the musical object as such, the latter coordinating perception and action. In consequence, the listener is projected inside the music by anchoring the gestures in its temporal unfolding, and in the process generating structural elements that shape experience.

### III. Situation, Depth, and Motor Intentionality

[3.1] Arguing that perceiving subjects engage in “dialogue” with the objects of their perception, Merleau-Ponty effectively shows that those subjects’ active bodies are the very source of how objects show up *as* objects. Our perception does not present us with inchoate, decontextualized shapes, colors, sounds, textures, and so forth, which our cognition then strings together into coherent forms. Instead, we see, hear, touch, taste, and smell things that are already meaningful to us—that have some value in our dealings with the world, because we *have* dealt with them as objects requiring some kind of bodily interaction. At stake here is the idea that musical relationships in the form of transformations—that is, “from within”—are not established by some metaphysical forces in the music and then mirrored, or extended, through gestures, but are constituted by different kinds of motor engagements on the part of listeners. These engagements are evidenced by recent behavioral and neuroimaging studies, which consistently show that listening is never passive, but that the motor system (both in the brain, as well as at the level of the entire body) is immersed in incessant activity, even when overt movements are explicitly suppressed.<sup>(20)</sup> However, as promising as these studies are for the burgeoning science of the embodied, enactive, and extended mind—science that endeavors to determine how bodily actions constitute cognition—what is conspicuously missing is the conceptual link that makes their findings relevant to music theory and analysis. Thus, in order to address the second question in §1.8 above (concerning the ways in which listeners’ bodies alter musical space and the constitution of its objects), in this section I will make the case for connecting musical experience with Merleau-Ponty’s phenomenology, focusing specifically on gesture as creating a particular kind of musical space, inhabited by particular kinds of musical objects.

[3.2] Studies have shown that gestures performed by listeners when engaged in musical activities are varied, idiosyncratic, and contingent on an assortment of conditions (Jensenius et al. 2010). These conditions include, among others, the listener’s socio-cultural background, musical expertise, genre and style of music, and sonic features that are particularly marked. They also include one’s corporeality and agility, both of which constrain the feasibility of certain actions. Yet despite their overwhelming variety, these gestures are not merely the result of a causal coupling between sounds and actions, but instead constitute a cultural practice of movement possibilities—they are not reflexes, but their specific shapes and temporal dynamics are guided by a kind of intentional comportment of the listeners’ bodies. Responding to the contingencies of the environment, gestures are “a negotiation between the given and the forged,” or between the world as “there,” and one whose meanings are generated in the act of gesturing (Noland 2009, 56).

#### *Spatiality of Situation*

[3.3] For Merleau-Ponty, gestures further constitute an embodied understanding that is the very condition that fundamentally makes cognition, language, and (ultimately) culture possible, a condition that unfolds against the background of what he calls an unconscious “spatiality of situation” ([1945] 2012, 102). We do not merely think ourselves in relation to an external world construed as a system of coordinates, each an objective point in space. Instead, we orient our bodies toward specific constraints and opportunities of the environment, something that readers familiar with J. J. Gibson’s (1977) theory of ecological perception will recognize as *affordances*.<sup>(21)</sup> Through this kind of orientation, and through motor habits coupled with different kinds of affordances, the world presents itself not as a set of determinate features, but as a network of possibilities framed by the perspective of our unique experience—as Merleau-Ponty asserts, “The body is our general means for having a world” ([1945] 2012, 147).

[3.4] The spatiality of situation as a possibility of language, cognition, and culture is thus for Merleau-Ponty a spatiality of the body. And as such, it emerges in actions, particularly in the way these actions inhabit space.<sup>(22)</sup> Again, this is not a Cartesian space of objective coordinates that stretch out around the subject, but a space in which the body’s situation lays down “the first coordinates,” and “anchors” the body in the objects that it faces in its various tasks. The placement of my body and the objects that elicit a skillful reaction create a unique perspective from which the world appears to me in a certain way. My actions are directed toward something, some object or activity, and it is in that object or activity that my body emerges. Nestled between the objective world and mental representations—both notions that Merleau-Ponty rejects with respect to perception—the body is “the third term” which forms the very condition of possibility for meaningful engagement with the world ([1945] 2012, 103).

[3.5] This framework, of course, seems to apply much more readily to performers, whose actions actually engage with physical objects, or even to dancers, whose movements are overt and available for scrutiny, than to listeners. Considering that



music cannot be grasped, or held, or shaped with one's hands, it seems that listeners are eternally destined to remain musical outsiders, capable only of indexing and measuring distances between musical elements with an intervallic yardstick.<sup>(23)</sup> As outsiders, they are separated from the world in a sense that events—for example, musical events that Lewin, after Husserl, calls “the determinable-X” ([1986] 2006, 60)—occur independently of their presence and unaltered by their actions. Such events are, indeed, pre-conceived, with the listeners' role relegated to passive submission to musical meaning. Music, meanwhile, turns into mere auditory input.

[3.6] However, according to Merleau-Ponty, the body is not passive. It is not fixed to its physical setting, but becomes embedded in what it attends to, in what it engages. “My body is wherever it has something to do,” he notes ([1945] 2012, 260); and again: “[The body is] a system of possible actions” ([1945] 2012, 260). The body itself is constituted in its poise, in its doing. Consider, for example, the following scenario. You observe your friend's four-year-old daughter trying to reach for a plate of cookies on a counter that is much taller than she is. Her attention is fully engaged with the plate, and she does not see that her hand is coming dangerously close to a glass jar sitting next to the plate. All of a sudden, her hand makes contact, the jar tips, and starts to fall over the edge of the counter. As a witness to these events, not only your attention, but also your bodily possibilities are fixed at the site of the falling jar. You saw that the child was about to tip it over, and you anticipated the fall, which is why you are now able to jump in and save the jar from shattering. In Merleau-Ponty's view, it is not as if your mind was running through plausible outcomes and calculating how much time it would take for you to get to the counter should your intervention be required, while your body remained passively grounded in its position. Rather, your body was poised and answering to the solicitations of the situation: it was anchored in the scene that was unfolding, and in which you were an active participant.

[3.7] Empirical evidence suggests that one of music's functions is to regulate listeners' bodily states, especially those involving group behavior: musical sounds engender coordinated movements in listeners, thereby also controlling affective dispositions of participants.<sup>(24)</sup> Indeed, listeners not only respond gesturally without reflection, but they often do so in ways that result in very similar reactions to the same sounds (Godoy 2010). Therefore, if we transpose the above view into our present discussion, listeners, through movements, become situated within the temporal processes that characterize the music they hear, and not in the space in which they perform their actions. In other words, *music* is the space of their situation. It solicits bodily involvement and beckons for one's motor actions, thus putting the embodied listener right inside the flow of sounds.

#### *Depth*

[3.8] This insideness of an embodied listener necessarily changes the topography of musical space, which addresses a concern expressed by Lewin, and offers an antidote to the kind of reification of musical objects that he scrutinized in his famous essay “Music Theory, Phenomenology, and Modes of Perception” (Lewin [1986] 2006). More precisely, he notes that it is the visual representation of music on a two-dimensional Cartesian plane, where noteheads pinpoint precise temporal and registral locations, that leads to impositions of “false dichotomies” with regard to musical events' ontology ([1986] 2006, 79). Geometric metaphors resulting from this representation induce claims that, given some system of classification, only one interpretation is possible for every musical event, and only one temporal flow persists in which these events may “exist.”

[3.9] Cartesian space and its geometry form the perspective of an omniscient being that surveys every element of the land at once, one that Merleau-Ponty—after the seventeenth-century Dutch polymath, Christian Huygens—evocatively calls *kosmotheoros*.<sup>(25)</sup> Such omniscience, which the French philosopher labels the “spatiality of position,” affects the very ontology of space, in particular the dimension of depth. Specifically, depth, rather than securing its own status as a unique presence in our experience, becomes equivalent to width viewed from a different angle. Look, for example, at a façade of a building. What you see is just that part of the structure, just the one side, but you do not *experience* the building as a flat surface. Instead, its depth is an element of how the building shows up in your perception, and this depth is unique in that it presents a set of possible actions that you can perform with regard to the building itself (e.g., you can enter it, you can throw a projectile inside through its windows, etc.). Now, where you saw depth in the distance between two objects, someone observing the very same scene as you, but situated off to the side, would see width. This is to say that your *experience* of the building's depth becomes their *observation* of its width—what was an action opportunity for you is just another measurable dimension (an interval) for them, a dimension precisely like any other. In this scenario, all unique features are flattened out, interchangeable, and devoid of markers relevant to experience. This “God-like” view is truly of an outsider who has no access to the objects that inhabit the space.<sup>(26)</sup> This outsider cannot interact with such objects; one can only observe them. It is precisely this space, whose origins Merleau-Ponty locates in the Renaissance technique of visual perspective, that denies the subject entry into the world of objects. They are instantly laid out before the subject on a geometrical grid devoid of subjectivity: depth as objectified and detached from experience (Merleau-Ponty [1945] 2012, 279).<sup>(27)</sup>

[3.10] For Merleau-Ponty, this detachment of depth from experience will not do, because depth is an “existential” dimension—it places the subject in the world, among objects. In other words, depth is what makes my experience of the world an experience of *me* in the world. In contrast to the objectified depth of *kosmotheoros*, the world of an insider has a “primordial depth,” something that links the subject with the objects of its world—something that places the subject *in* the world. In “Eye and Mind,” his essay on painting, Merleau-Ponty notes that depth is “a voluminosity we express in a word when we say a thing is *there*” (1964, 185) For him, it is something through which things are seen, as well as that which is

hidden but which exerts an ontological presence on what it is that we do see. It is not a measurable dimension in itself, one that can be grasped and quantified; rather, it is that which gives all other experienced dimensions their voluminosity. Merleau-Ponty calls this an “openness upon the world,” whereby depth is constitutive of the experienced dimensionality of our environment.<sup>(28)</sup>

[3.11] Returning to Lewin’s concern, musical objects in two-dimensional representations, despite being given unique names based on their spatial location, are indistinguishable from one another. Or rather, they are only distinguishable by name and location: like on a map, they are flattened out and speciously systematized. They have no depth in an existential sense. However, if we reconceive it through the lens of the spatiality of situation, (musical) space is no longer a Cartesian plane, and (musical) distances can no longer be measured with intervals. Consequently, the solution offered here is to give depth to each musical event by regarding it as already pregnant with meaning that is not audibly present to the listener, but is nevertheless a fundamental component of experience. When we realize it in this way, every present sonic phenomenon is imbued with its absence, with the background against which I, as a listener, bring it into its singular being.<sup>(29)</sup>

#### *Motor Intentionality*

[3.12] Instead of our sensory organs equally absorbing all incoming stimuli, only a sliver of the world is available to us at a time. We need to move and interact with it in order to apprehend it, and we need depth as a background against which to understand objects that make up our environment as meaningful, holistic Gestalts. Unlike Descartes’s *cogito*, we are not detached from our world, but belong among objects. At the same time, we ourselves are *not* collections of objects. Rather, we exhibit a “comprehensive body *purpose*” in what Merleau-Ponty calls our body schema (*schéma corporel*) (Merleau-Ponty [1945] 2012, 101). Construed as an “integrated set of skills poised and ready to incorporate a world” (Carman 1999, 219), this schema unites our bodies and allows depth to emerge as an existential dimension. To put it somewhat simplistically, it makes the body “whole.” In modern cognitive terms we would say that it is a dynamic, ongoing process of monitoring the body’s position in three-dimensional space, a process that unfolds below our level of consciousness, and involves “certain motor capacities, abilities, and habits that both enable and constrain movement” (Gallagher 2005, 24). It is what gives an expert violinist, for example, the ability to play the instrument without consciously tracking and controlling every movement of every finger in contact with the fingerboard, the angle and pressure of the right hand on the bow, balance and breathing, and so on. But it is also the very mechanism that allows humans in general to skillfully engage with their world, in that we are all “experts” when it comes to handling various objects and terrains in order to achieve desired outcomes.<sup>(30)</sup> Thus, the body schema is something that underpins *all* the actions that unfold with proficiency and intent, whether we are dexterously executing three-octave double-stop arpeggios on the violin, or simply picking up our favorite mug to take a sip of coffee.

[3.13] Unified by the body schema in the context of its spatiality of situation, the body projects into its environment what Merleau-Ponty calls a pre-conceptual, pre-cognitive *motor intentionality*. Based on Husserl’s ideas that consciousness is intentional because it is “about something”—which we saw above with regard to Rings’s interpretation of Lewin’s “doing”—motor intentionality refers to movements that are “about something” external, something that solicits an action, and involve a bodily and situational understanding of space, its features, and the objects that occupy it. A grasping movement, for example, is motor intentional because the completion of the movement is already present at the moment of its initiation: the object’s properties—its size, shape, weight, density, content, and so forth—can be identified by the shape and formation of the hand. At the same time, motor intentional movements are not foreordained; they remain indeterminate until the moment of their fulfillment, because they are responsive to the dynamics of the world.<sup>(31)</sup> They are also what Merleau-Ponty calls “pre-predicative,” because they do not offer the kind of information about the object that can necessarily be described using language. In this view, actions performed skillfully, and so exhibiting motor intentionality, constitute behavioral phenomena that fall somewhere between mechanical reflexes and full-blown cognitive processes (Kelly 2002, 167).

[3.14] Kelly further comments that “motor intentional activity . . . essentially discloses the world to us . . . but cannot be captured in the process of doing so” (2002, 389). Although it can be reflected upon in the aftermath, at the moment of performing an action there is a bodily understanding that is logically different from a cognitive kind of comprehension. In motor intentionality, there is no way to represent a feature of the world apart from actually performing some action that is directed toward that feature. Put differently, the content of a motor intentional action, and the attitude toward that content, are indistinguishable, and such an action “is directed not just toward a location, but toward a located object” (2002, 384).

[3.15] Such a definition felicitously applies to gestures performed by listeners. While they do not bring the gesturer in contact with physical objects—unlike movements of instrumentalists—they are nevertheless motor intentional, because they respond and conform to sonic features of music in predictable ways. Based on empirical observations of moving listeners, researchers have shown that, for example, loud sounds elicit large movements, often with high velocity and acceleration; legato sounds, or sounds with gradual attack and decay envelopes, engender movements that are fluid (with a high ratio of velocity to acceleration); short sounds result in gestures that end abruptly.<sup>(32)</sup> More examples abound, in each case illustrating predictable and intersubjectively verified correlations between musical events and gestural features, attesting to the motor intentionality that underpins listeners’ gestures. Crucially, if we accept Merleau-Ponty’s stance regarding situation, depth, and motor intentionality, these listeners are uniquely positioned to reconfigure musical space, and to reveal elements of musical structure that an “omniscient” perspective cannot. Not only does each listener constitute “objects” of experience anew, but

these very objects show up as Gestalts with different kinds of depth, depending on the bodily capabilities of the gesturer.<sup>(33)</sup>

[3.16] By calling on Merleau-Ponty's phenomenological encounter with an active perceiver we can sketch the following provisional picture of characteristic gestures that will allow listeners to directly participate in transforming sounds into interrelated musical objects. To begin, by having a motor intentional disposition toward musical sounds, listeners constitute those sounds as entities that require some kind of action that is constrained by each listener's background, motivations, physical capabilities, and so forth. Such constitution effectively turns sounds into musical objects, rather than these objects being given in advance. Meanwhile, the action itself creates depth in the musical object as it becomes constituted: it crystalizes a sonic context that imbues the object with a specific, unique meaning for that listener, as well as the very ground of experience of that object. Finally, the listener's particular spatiality of situation from which such depth emerges places the motor-intentional actions of that listener inside the music, thereby generating further musical objects in an ongoing process of attending to a piece of music.

[3.17] While admittedly tentative and fragmentary in its current form, such a picture already furnishes us with a framework within which listeners can join Lewin's idealized dancers and performers in actively recasting the musical soundscape from one of spatial intervals into experiential transformations. Continuing with this model, in the following section I discuss some important consequences of this reconfiguration, and present a case for using gestures as a ground for analytic engagements with music.

#### IV. Gestures in Analysis

[4.1] Placing active listeners inside music draws attention to the dynamics of depth and situation as constitutive of singular musical objects, thereby creating the conditions for using those listeners' bodies as foundations for analytical understanding. In what follows, I will suggest one answer to the third question posed in §1.8 above—concerning the practical aspects of an embodied music analysis—by arguing that such foundations can be turned into models of musical processes that can further ground more traditional, narrative expositions of musical understanding. However, before reaching that point, musical gestures of listeners need to overcome a crucial obstacle. Namely, they seem in practice to fall into a kind of intellectual no-man's land, adrift between the rigorous logical operations of formalism and the speculative elasticity of an indeterminate body. Finding oneself within it, one risks becoming too singular, too introspective, and too self-indulgent for the hope of intersubjectivity and what Lochhead (2006) calls "sharability." On the one hand, musical gestures often lack specificity with regard to their referent. Pointing to an under-determined signified, they can quickly dissolve into something like an interpretive dance: artful and aesthetically pleasing, to be sure, but not very effective at elucidating an analytic understanding of the musical object with which they are coupled. On the other hand, they appear far too idiosyncratic and particular to be useful for music theory and analysis. Every gestural articulation is different from the next, its structural and expressive morphology just as fleeting as that of the music that engenders it. Thus, to talk about a *real* musical body—not one constructed and idealized by a theory, but the moving form of an actual listener—seem equally unfeasible, profoundly self-absorbed, and decidedly *un*-sharable. Despite the visible, interpersonal trajectories of human movement, the semiotic imprecision and the morphic uniqueness of the sheer variety of muscular efforts—far greater than other communicative signs available to us, including language<sup>(34)</sup>—seems to guard against their incorporation into a more systematic analytical endeavor.

[4.2] Yet the body considered as an abstract theoretical category—while capable of sustaining discourse that is in some ways more universalizing and intersubjective—is also problematic, because it loses its very purpose relative to the task of analysis at hand. Paradoxically, it becomes as passive in its reception of musical sounds as the computational, disembodied mind that it is supposed to augment, or replace altogether: it doesn't *do* much in terms of actual *doing*. Zbikowski (2013) locates one example of the body's ossified condition in the writings of Jean-Luc Nancy, particularly his monograph *Listening*. According to the former, Nancy excludes the body from representations of musical behavior, "turning it into an empty shell into which can be poured the passions inspired by sonic phantasmagoria" (2013, 106). Rather than a vehicle for active engagement with sound, the abstract body is actually an immobilized product of its own socio-cultural forces that predetermine its behavioral functions, a situation that Zbikowski also detects in contemporary musicology. A body like that does not produce its own condition, or constitute its own objects of perception—those aspects of existence are given in advance, and this is precisely the state lamented by Lewin.

[4.3] In the final section of his "Phenomenology" essay Lewin famously dismantles the very foundation of his prior analysis of Schubert's "Morgengruss" by claiming that it is "dangerous" to assume "that music theories are, or should be, fundamentally perceptual in nature or purpose" (Lewin [1986] 2006, 94).<sup>(35)</sup> The claim is supported by an assertion that the perceiving listener finds music "given and there, not just sensible and present." In other words, music for the listener comes pre-constituted: the activity of perception in this account is passive and accomplished *post hoc*. Meanwhile, what Lewin desires is for music theory "to be useful beyond analysis and perception as goads to musical action, ways of suggesting what *might* be done, beyond ways of regarding what *has* been done" ([1986] 2006, 96). The roles of composers and performers both fit this bill, because here the relevant persona produces something new, something that did not exist prior to the activity of composing or performing. Crucial in Lewin's assessment is the distinction between action and perception—which he equates with doing and understanding, respectively ([1986] 2006)—and it is precisely this distinction that stops him from supporting



music perception as a basis of music theory. Music theory for him seems to require a generative impetus beyond mere prescriptions, in that it must allow for the possibility to regard musical processes and structures as continually renewable, and create the conditions to make this renewal attainable. On the Husserlian view that Lewin endorses, perception does not, in fact, fulfill that role.

[4.4] However, if we continue to draw on Merleau-Ponty, we need not jettison this generative impetus together with Husserl's phenomenology. In fact, for the French philosopher perception not only involves interacting with the world, but it is through this interaction that things we perceive are constituted. In other words, perception is an active process that generates objects, not a passive engagement in which phenomena are merely presented to a disembodied consciousness.<sup>(36)</sup> Unlike Lewin's view, here perception is emphatically *not* apperception, meaning that it is not grounded in previously acquired representations of the world, on the basis of which currently experienced events can be understood. Instead, perception *is* the behavior—gestures, actions, stances, postures, affects, and so on—that shapes experience. To view real, touchable, singular bodies in action is therefore to witness the creative, animate power of perception. In short, *perception* is something you *do*.<sup>(37)</sup>

[4.5] To balance the critiques enumerated above we can consider the way in which the body synthesizes subjective experience while making it available for external evaluation by a third-person observer. In his discussion of human spatiality and motility, Merleau-Ponty makes the following observation: “Within the busy world in which concrete movement unfolds, abstract movement hollows out a *zone of reflection and of subjectivity*, it superimposes a virtual or human space over physical space” ([1945] 2012, 114; emphasis added). Unlike concrete movements, which involve either actions performed to accomplish a specific task (e.g., brushing one's teeth, fixing a carburetor, or playing the piano) or a literal reenactment of those movements, abstract movements are performed for their own sake. This is to say that they are not characterized by a kind of intentionality that is directed toward manipulating physical objects. Such gestures are exceptional, because they demonstrate that we are not necessarily causally tied to the world in which we exist. Although our bodily presence provides a background of possibility for perceiving that world, thereby acting as a vehicle for movement, the same body can become an *end* of movement. As Merleau-Ponty puts it, we are capable of “breaking with [our movements] insertion in the given world” ([1945] 2012, 114).

[4.6] Gestures that listeners perform in response to musical sounds realize precisely this breaking-with, in that they are not carried out with the intention of affecting the physical makeup of the world. Rather, they secure a human space of reflection and subjectivity without pointing to anything physical in particular. Guided by the subjects' pre-conceptual understanding of music's unfolding processes, and drawing on those processes for cues concerning how to constitute subjective musical space and time, these gestures are absolved of the constraints of musical instruments and can be executed in ways that defy the laws of sound physics. Meanwhile, the subjectivity and reflectivity of cognitive and perceptual mechanisms that underlie their creation are balanced by the visual and kinesthetic presence of the body. To gesture is to demonstrate—to engage others with a physicality that is remarkably interpersonal and inter-corporeal.<sup>(38)</sup>

[4.7] Still, in practice it seems exceedingly arduous to get away from introspection and self-indulgence when considering an analysis that begins with the subjective body of the listener, any more than it is to generalize from first-person accounts of performers. As Nattiez writes, “no analysis is truly rigorous unless written down . . . since the *record* of the analysis enables it to be checked: once it is written down it is possible to review, criticize, and go beyond an analysis” (1982, 244; emphasis in the original). In other words, analysis needs to participate in the institutional accumulation of “knowledge,” understood here strictly as that which is written. As Nattiez claims, a review or critique of an analysis rests on the reviewer's ability to retain the information presented by the analyst, but the task is imperiled when all that the reviewer has access to are the listener/analyst's fleeting and semiotically obscure gestures.

[4.8] Although Nattiez's position is merited, we must note that analyses—even when presented as narrative descriptions, and intended to convey the analyst's interpretations in a way that needs to be followed linearly—more often than not also include various kinds of models and visual representations. Unlike prose accounts, these are holistic snapshots of the analyst's reductive, deductive, and inductive processes. They neither represent the music, nor create verbal reports, but instead concretize an analytical perspective that may include activities like extraction, recomposition, recontextualization, abstraction, graphing, and, above all else, depiction. These models exist as something between a lettered testimony and an embodied understanding of music's processes: they mediate the phenomenal experience of music and a linguistic record of that experience.

[4.9] To be clear, I am not suggesting that the gesturing body itself mediates between musical sounds and an analytical understanding; such a view would take us right back into the territory of quasi-embodiment and Cartesian dualism.<sup>(39)</sup> Quite the opposite, my claim is that by projecting a motor intentionality toward musical events, the body is *already* doing analytical work, and various ways of quantifying and visualizing that work can be useful in creating verbal accounts of musical structure. As Zbikowski—embracing psychological research on the various functions of gestures in communication—points out, “gesture offers a dynamic, imagistic resource for conveying thoughts that would be cumbersome to express through language” (2011, 89). Indeed, in contrast to analyses grounded in performers' actions that I presented in the introduction to this article, Zbikowski elsewhere (2008) draws on known dance patterns to show how certain (in this case, eighteenth-century) musical grammars have emerged in response to bodies' movements in space. He then uses those patterns

—specifically, the steps to a *pas de bourrée*—as a basis for his analysis of the Finale from Haydn’s String Quartet op. 76, no. 4.

[4.10] Bringing in the listener’s body, however, need not be limited to gestures and actions that are choreographed or otherwise codified, and can be equally illuminating in less systematized contexts. Given the foregoing discussion, what I propose is, therefore, to build analytical models by drawing on all sorts of bodily actions performed in response to musical sounds; to use the body’s motility as a source of musical knowledge by correlating various quantitative features of gestures—such as acceleration, velocity, and relative orientation in space—with musical sounds, and to visualize those correlated features in order to draw out analytical narratives. Doing so not only allows us to consider complex sonic elements, but also to consider them from a phenomenologically relevant perspective, one that affirms the centrality of situatedness and depth to experience. That is, it lets us balance between subjectivity and reflection of the listener on the one hand, and the body as a passive figure seen from a distance on the other, thus preventing either perspective from overwhelming the discourse and jeopardizing sharability and claims of veridicality.

## V. Analytical Snapshots From the Body

[5.1] We are now ready to collect the answers I suggested to the three questions posed in §1.8, and apply them to analytical examples by examining how the spatiality of listeners’ bodies participates in the formation of musical structure. To recall, for Merleau-Ponty ([1945] 2012, 105) this spatiality—that is, the manner in which we inhabit space and time—emerges in and through our actions, and we can come to understand it only through the study of our movements. Accordingly, to engage with the body in music analysis we must have a grasp of how specific gestures correlate with musical sounds—we need to observe real bodies, in all of their motile, corporeal voluminosity, all of their vulnerability and frailty, all of their insurgency and impudence. Thanks to recent technological advances, we are able to do this in methodologically robust ways by incorporating motion-capture tools into the analytical process. All motion features presented below were obtained in exactly this way, allowing us to examine both individuals and groups of participants. In brief, listeners—both with and without musical and dance training (hereafter “musicians,” “non-musicians,” and “dancers,” respectively)—were asked to move freely and without restrictions to nine 30-second excerpts taken from recent Western instrumental art music repertory.<sup>(40)</sup> Outfitted with reflective markers, their movements were recorded with infrared cameras, and all resultant position data were turned into acceleration profiles. Although free movements in space are eminently diverse, exhibiting more or less subtle variations and nuance in many different spatio-temporal ranges, previous research has shown that movement acceleration is best correlated with discrete musical events, which is why this particular dimension was used as a basis for the foregoing discussion.<sup>(41)</sup>

[5.2] As concerns the present study, three musical fragments—taken from Olga Neuwirth’s *Vampyrotheone*, Elliott Carter’s *ASKO Concerto*, and Thomas Adès’s *Living Toys*—will illustrate different ways of using motion-capture data to make models drawn from bodily engagements with music. One characteristic that these pieces share is that they present rhythmically and timbrally complex passages that engender very particular gestural responses. I will use the analyses to show, respectively: (1) how sounds that are somehow underdetermined, but which feature very rich timbres, can be subsumed under a single, unifying gesture; (2) how acceleration profiles can be used to discover structurally important musical features that show up in listeners’ experience; and (3) how these profiles can reveal aspects of movement that constitute a meaningful element of listeners’ experience, but that are not correlated with anything literally present in the auditory signal itself.

*Neuwirth*

[5.3] The Austrian composer Olga Neuwirth’s *Vampyrotheone* for three soloists and three chamber ensembles (completed in 1995) offers the listener few opportunities to attend to exact pitches, unambiguous rhythms, and recognizable timbres. We might regard it as a textbook example of what David Metzger (2009) calls the *sonic flux*—taking sound in its totality as a vehicle of musical meaning—where the auditory experience teeters somewhere between noise and “musical sound.” The piece opens with a low rumble on the lowest string of the piano, played tremolo with soft mallets, that quickly builds to a *fff* chord immediately following the downbeat of m. 2. The buildup is effected with a rapid crescendo, as well as through the addition of other instruments that, either stepwise or with indeterminate glissandi, ascend in pitch. A triplet-eighth after the downbeat of m. 2 everyone simultaneously attacks a dense, complex sonority (its pitch content is shown in **Example 1**), then gradually backs off from the climax and eventually gives way to a softer background.

[5.4] Predictably, the *tutti* attack in m. 2 engenders a clear, unambiguous gestural response, which can be seen in the top panel of **Example 2**. Note specifically the sharp peaks in acceleration around 500 milliseconds. Indeed, this is the most obviously articulated moment in this excerpt. Together with the bongos, trumpet, and piano simultaneity on the second triplet-eighth-note after the downbeat of m. 6 (corresponding to the peaks around 2500 milliseconds), the two timepoints bookend a far more opaque, amorphous sonority unfolding in mm. 3–5. It is this sonority that I want to discuss here.

[5.5] Sounds in these three measures might best be characterized as “scraping,” “grating,” “metallic,” or any other adjective in this family. The most aurally distinctive—and at the same time the most abrasive—sound is that of the cello playing between the bridge and the tailpiece, taken over by bass clarinet and baritone saxophone multiphonics at the end of m. 4. A transition of sorts between these two events is effected by cowbell tremolos. There are also other instruments that come into and out of the texture, including a piano tremolo in m. 4, electric guitar dyads in mm. 3 and 4, and perhaps most notably, the

undulating, pitchless breaths of the bass trombone, tuba, and horn. To give this moment a poetic tinge, imagine an aqueous musical surface strewn with sonic ripples that fold into each other.

[5.6] While the sonority may be complex in terms of its unfolding timbre, it has a kind of essence that binds it into a surprisingly coherent unit. This binding is entirely contextual (how often does one construe “good continuation” as a succession of cello noise, cowbell, and woodwind multiphonics?), but its effect is reflected in participants’ acceleration profiles (shown in Example 2), whose low values between mm. 3–5, represented on the vertical axis, are consistent with relatively smooth, unchanging gestures. Indeed, the lack of clear articulations, regular time markers, and discrete pitches tends to engender fluid actions that are themselves largely unarticulated. By matching the temporal and timbral dimensions of sounds, we can regard these actions as constituting complex musical objects that might otherwise be difficult to classify, or even to describe. Performing and observing these gestures turns into a display of an embodied understanding of music’s dynamical processes. Because such a display synthesizes not only all of the disparate sounds heard in this passage, but also all of the different performance actions that go into producing those sounds, the body becomes a nexus of sometimes competing musical meanings and expressions. And as it does, it acutely problematizes theories based on the “listener’s body mirrors the performer’s body” paradigm described in the Introduction.

[5.7] Indeed, the way in which gestures absorb the multitude of sonic dimensions highlights a possibility that the sound in mm. 3–5, the one which seems so eager to resist clear definition, categorization, or even basic description, might play a structural role. Namely, the whole piece is made up of short episodes that are unified by timbral qualities, and juxtaposed without any hint of an intelligible narrative or inherently motivated progressions. As soon as a process begins to unfold, Neuwirth shuts it down and starts a new one. This reluctance to present a coherent shape brings to attention the listener’s body, specifically its role in subsuming, or folding in, disparate sounds into actionable wholes. This is to say that the expressive character of these sounds is made intelligible through their envelopment within bodily gestures. They are given a temporal structure and become imbued with affective potential, instead of simply being strung up in succession without a coherent trajectory. Indeed, gestures unify disparate sounds—sounds that exhibit altogether different timbral characteristics, rhythmic profiles, dynamic shapes, and so on—into episodes. As a result, different units, even ones that are non-contiguous, can form relationships to each other by engendering similar bodily responses in listeners.

#### *Carter*

[5.8] The manner in which Neuwirth juxtaposes different sonic effects places musical events in sharp relief, in turn marking them as unambiguous affordances for movement. As a result, nearly all participants moved in the same timeframes. Still, many contemporary pieces express affordances that are a lot more evenly distributed in time, resulting in far less intersubjective agreement. To put it differently, when presented with music in which contrast is attenuated, listeners do not necessarily focus on the same events. This is the case with the opening trio (mm. 20–54) in Carter’s *ASKO Concerto* (2000), a fragment that features a characteristic compositional technique in which three separate melodic lines are presented at once, each with its own unique pulse. Due to the resultant rhythmic complexity, an examination of all movement profiles reveals little agreement as regards acceleration values, indicating that participants moved with varying effort and energy.<sup>(42)</sup>

[5.9] However, this does not mean that gestures are altogether meaningless insofar as their role in indicating to the observer which events in the music are experientially salient and potentially structural. Instead, it suggests that gestural interpretations of the musical surface are a lot more idiosyncratic, and vary considerably from one participant to the next. In such cases we can look at the total number of acceleration peaks, instead of specific acceleration values, in each timeframe. This signals the level of agreement among participants with regard to the temporal placement of gestures—*when* they moved—which can be represented as a histogram. Shown in **Example 3**, the vertical axis indicates the number of participants whose gestures created a peak in acceleration for every 350 milliseconds timeframe, displayed as a percentage of the total number of participants. For example, at point (*a*) at 6 seconds, about 65% of participants (10 out of 15) moved in a way that resulted in a peak in acceleration, regardless of how high this peak was relative to their remaining gestures.

[5.10] Remarkably, when expressed in this format, we do see the participants gesturing in response to the same sonic events. Especially telling are timeframes in which agreement reaches over 50%, or when more than half of the listeners moved at the same time. In addition to the moment (*a*) just discussed, further along we observe three prominent peaks between 15 and 20 seconds, the highest reaching 80% (labeled *b*), which indicates that 12 out of 15 participants moved at precisely this moment. There are also smaller peaks of 60% toward the end of the excerpt. Altogether, this information suggests that at these timepoints something special is happening in the flow of musical sounds, inviting us as analysts to examine more closely what yielded such agreement.

[5.11] Take, for example, the peak (*a*) at 6 seconds, which corresponds to a single, seemingly unexceptional descending dyad in the oboe (D $\sharp$ 5–G $\sharp$ 4 in m. 23; see **Example 4** for a score reduction). Because the melodic strand of each instrument is fragmented, it is likely that throughout the entire excerpt listeners were generally tuning into the overall, holistic effect of the combination of the three instruments. However, this dyad is singled out: not only is it exposed in terms of timbre and register, but it also appears on its own, following a continuously shifting compound melody which involves the participation of all members of the trio (mm. 20–23). In fact, this isolation can be taken as a structurally meaningful moment, and its

relationship to gesture may mark it as a unifying process in this section of the *Concerto*.

[5.12] Consider first the interval of a descending perfect fifth. Although it is an inversion of the ascending fourth that opens this trio (see oboe in m. 20), it is rather unlikely that listeners were immediately aware of this relationship. Instead, the turnaround itself—the fact that the music changed directions from an ascent to a descent—might have played a significant role in shaping those listeners’ experience. As **Example 5** shows, the first four measures of the fragment are characterized by rising dyads, many of which are articulated *legato* with slight emphasis on the first pitch in each pair. Thus, the specific gesture that would correspond to these dyads is established right away as a relevant aspect of the piece’s organization. Looking ahead, we can see that the oboe turnaround in m. 23 initiates a reversal in the direction of the dyad, while at the same time maintaining its articulation and (for the most part) emphasis. From that point, nearly every measure until the entrance of the “ritornello” in m. 55, which marks the end of the trio, contains at least one such dyad (shown in **Example 6**).

[5.13] Abstracting from the dyad’s gestural correlate, Carter actually deploys the dyads in a way that gives this fragment coherence by making use of their musical properties. An important technique introduced in the first four measures, and which becomes much more characteristic in the remainder of the trio, is the use of exact pitch repetitions. These repetitions result in a sense of continuity from one dyad to the next, and between dyads separated by longer spans of time. Note, for example, the dotted arrows in Example 5, which show precisely these connections. This technique gains prominence in the latter part of the trio, as illustrated by the solid black boxes in Example 6. Especially salient in that regard are mm. 47–53, which feature a running common-tone line (see **Example 7**). This line participates in a buildup of intensity, amplified by the rising register of those common tones (from E $\flat$ 4 in the viola in m. 47 to C6 in the oboe in m. 53). Two more places where this technique appears are mm. 32–34, which feature a playful exchange between the viola and the oboe, and mm. 37–41, where the oboe and the horn engage in trading descending half-steps.

[5.14] Carter applies the same principle of common-tone continuity to the four longer descending gestures scattered throughout this section of the *Concerto*, including the last two utterances (see oboe and viola in m. 54). Compare, for example, the oboe gesture in m. 28 with the one in m. 54, or the oboe descent in m. 36 with the viola in m. 54. Even though there are no significant pitch-class-set associations between them, each pair of descents begins with the same pitch. Not only that, each pair retains one of the intervals from one iteration to the next: the common interval between statements in mm. 28 and 54 (oboe) is 4 semitones, while between mm. 36 and 54 (viola) it is 6.

[5.15] To be sure, I am not claiming that listeners whose movement profiles were used in modeling the above analysis somehow sensed that the descending dyad in m. 23 was about to become an important element of the subsequent texture. I cannot even suggest that this moment was marked as logically coherent in their experience, or that, when questioned, they would have been able to explicitly articulate what was happening in the music. However, the way Carter presents this dyad *did* elicit intersubjective agreement with regard to the temporal placement of gestures, and when considered within the theoretical framework presented earlier in this article, this agreement is indicative of an intersubjective motor-intentional understanding of the work’s experiential structure. In turn, the amalgam of these gestures, visualized as a histogram, directed me toward an interesting process that pulls together this section of the *Concerto*.

#### *Adès*

[5.16] Motion-capture observations typically reveal a close relationship between movement acceleration and sound events, consistent with the notion that listeners respond to what is actually present in the auditory signal. Indeed, in some instances that relationship is remarkably robust, and can be used to predict which sonic features will elicit gestural responses of a particular kind. As I already mentioned above, loud sounds, for example, result in big gestures with high acceleration, which was exactly a response that we saw in *Vampyrotheone*. There are times, however, when our observations disclose movements that do not seem to respond to anything literally *in* the sound itself, but which nonetheless form a salient element of listeners’ experiences. Such is the case with several participants who gestured to Thomas Adès’s “Militiamen,” the fourth movement from his chamber work *Living Toys* (1993).

[5.17] As the title suggests, the movement conveys a military affect, accomplished through the use of a snare drum and trumpet duet in the opening measures.<sup>(43)</sup> Such a combination is typically associated with unwavering regularity, where the two instruments are meant to instill in marching soldiers a sense of pulse, and to engender among them what William McNeill (1995) calls “muscular bonding.” However, Adès’s version presents only a grotesque, twisted remnant of that soundscape. First of all, the trumpet plays in its own metrical realm, and is explicitly directed to sound as if the part were improvised, thus eschewing coordination with the snare drum. Secondly, the snare drum works hard to obscure any feeling of a pulse, let alone meter, by rapidly switching between duple and triple subdivisions of eighth-notes, and consistently avoiding articulations on the dotted-quarter *tactus* of nearly every measure (see **Example 8**). Finally, the percussion meter itself is a compound triple ( $\frac{3}{8}$ ) which does not square with the military affect at all, where duple meter provides maximum efficiency because it allows the symmetry of the marching body (“left, right, left, right . . .”) to map onto patterned sounds. In fact, meter in “Militiamen” in general is attenuated, since among the percussion instruments only the field- and bass drums articulate beats 1 (with an eighth-note pickup) and 3.

[5.18] It is, therefore, quite curious that some participants’ acceleration profiles demonstrate the presence of a beat that is

actually absent at the musical surface. An especially interesting case concerns the participant whose profile is shown in **Example 9**. Even a cursory visual assessment indicates that there is a regularity throughout the length of the entire recording, a finding that is confirmed by a more detailed, quantitative examination. After adjusting for various micro-timing variations, the average time between peaks in acceleration is 350 milliseconds, and it remains quite steady for about 7 seconds (from 4.5 seconds to 11.5 seconds), which corresponds to mm. 301–4. Indeed, the standard deviation from the mean here is only 23.5 milliseconds, suggesting that the participant's body generated some kind of steady resonance frequency. What is downright uncanny is that this 350 milliseconds period is almost exactly the eighth-note pulse (357 milliseconds, or eighth note=168MM), if such a pulse were actually articulated in the auditory signal!

[5.19] Of course, we cannot be absolutely sure that this participant (and others whose acceleration profiles exhibit similar properties) somehow derived such a pulse from what they heard. What is far more likely is that their body's dynamics fell into a stable frequency that only happens to *coincide* with the absent eighth-note period.<sup>(44)</sup> However, what *is* remarkable is that the finding points to a non-existent musical feature that nevertheless structures experience. One plausible explanation for this correlation is that regularity was insinuated—leading to the listener's expectation and behavior—by the particular instrumental combination. Even if, hypothetically, the period of their movements did not match the implied beat, its presence and consistency with which it appears would have offered an interesting window into their embodied understanding of what was happening in the music. Importantly, this observation suggests that some musical objects are constituted *only* in behavior, and have no “determinable-X” kind of existence in the music itself (if we take “music” to refer solely to the auditory signal, which is a problematic assessment to begin with). To what extent we want to incorporate such objects into analytical narratives is, of course, a matter of choice regarding the kinds of insights we want to share with interested parties. However, their existence already creates an opportunity for serious conversations about the role of listeners' bodies in musical experience, and how much of that role we wish to foreground.

## VI. Conclusion

[6.1] Listeners' gestures are multidimensional. They are multidimensional in a traditional sense, in that they take up the dimensions of space and time, but they also have energy and effort, all of which can be quantified and represented graphically. As I have shown above, this multidimensionality of actions can be an advantage for analyzing the multidimensionality of music, even when considering a single quantifiable dimension: acceleration. For example, it captures within a single trajectory musical processes that unfold on different temporal levels, processes that exhibit such complex dynamics of unfolding that cannot be modeled with a single operation, or even processes that can *only* be captured by the intentionality of actions. In other words, the gestural body subsumes individual components of a dynamically ongoing event into a holistic Gestalt, distinct but indissoluble.<sup>(45)</sup> Of course, while in the interest of space I only focused on acceleration, other (no less critical) movement features—such as fluidity, spatial orientation, and what we may loosely call “figurative meaning”<sup>(46)</sup>—can also be used as models for analytical inquiry by revealing other structurally important elements of music.

[6.2] In this article I have shown how analysis from listeners' bodies can be accomplished by using those bodies as models for musical processes, and created a theoretical framework in order to contextualize those bodies' capabilities within the broader concerns of phenomenology, experience, and music analysis. Rather than actions contingent on the behavior of performers, or responding “automatically” to isolated sonic features, listeners' gestures are actually a category of meaningful, motor intentional movements. By eliminating the causal link between objectively located musical structures and their purported experiential effects, my approach also highlights the contingency of musical understanding by emphasizing what Voeglin describes as “the particularity of the listening subject in the contingency of his experience” (2010, 14). Instead of revealing meaning inherent in combinations of sounds, analyzing from the active body of the listener accounts for the possibility of generating, amplifying, and altogether transforming meaning with each listening. At the same time, it does so in ways that can overcome the idiosyncratic nature of listeners' gestures by turning them into visible, motor intentional models of musical sounds.

---

## Appendix

[A1] The purpose of this observation study was to explore how listening participants respond gesturally to complex excerpts from contemporary instrumental art music. Of particular interest was their gestural rendering of music without a strong sense of meter or pulse. Also pertinent to the present article was the fact that listeners did not receive visual cues regarding the movements of performers.

### *Selection of Excerpts*

[A2] 29 self-identified expert musicians evaluated 20 excerpts, each taken from twentieth- and twenty-first-century Western instrumental art music and lasting about thirty seconds, on the basis of the following criteria: (1) the absence of any perceivable pulse (“no pulse”); (2) the presence of a pulse but no metrical organization (“pulse, no meter”); and (3) metrical organization (“meter”). Each excerpt was preceded by a 500 millisecond sine-wave tone (880Hz) and one-second silence to alert the participants that music is about to begin. A single tone, instead of a number of regular ones, was used to prevent participants from entraining to an auditory signal before the music started.



[A3] Based on musicians' ratings, the following nine excerpts were chosen (three with the highest ratings in each category):

(1) No pulse

Excerpt 1: Olga Neuwirth, *Vampyrotheone*

Excerpt 2: Philippe Fènelon, *Diagonal*

Excerpt 3: Toru Takemitsu, *Coral Island*

(2) Pulse, no meter

Excerpt 4: Pierre Jodlowski, *Barbarismes*

Excerpt 5: Elliott Carter, *ASKO Concerto*

Excerpt 6: Olga Neuwirth, "The Long Rain"

(3) Meter

Excerpt 7: Harrison Birtwistle, *Exody* (a)

Excerpt 8: Harrison Birtwistle, *Exody* (b)

Excerpt 9: Thomas Adès, *Living Toys*, "Militiamen"

#### *Participants and Task*

[A4] 44 volunteers—all of them right-handed—participated in the motion capture experiment. The majority of them were undergraduate and graduate students, and faculty and staff at the University of Oslo in Oslo, Norway. Based on a modified Ollen Musical Sophistication Index questionnaire, they were divided into three groups: musicians (15 participants), dancers (13 participants), and non-experts (16 participants). Modifications to the questionnaire involved the inclusion of questions about dance experience.

[A5] Each participant was tested individually. All participants first listened to the nine excerpts to familiarize themselves with the music. They were then told that the task of the experiment was to move along with what they heard. We purposely did not constrain their movements in any way, letting them determine on their own how to interpret the directions of the task, and allowing complete freedom of movement. After participants were fitted with a motion capture suit; we explained that they could use their whole bodies and move unrestricted within the capture space, but that we were especially interested in movements of their right arms. Each excerpt was then played three times in a row: the first time for participants to be able to listen to the excerpt with minimal distractions, and two more times so that they could move in response to the music.

#### *Motion Capture*

[A6] Four reflective markers were placed on the right (dominant) arm and hand of each participant in the following positions: wrist, elbow, shoulder, and the seventh cervical vertebra. Movements were recorded using Qualisys infrared motion capture cameras at the rate of 100Hz. The capture space for the system was calibrated to include an area of 3 meters x 3 meters, from the floor to about three meters in height, thus allowing comfortable and relatively unrestricted movements of the whole body.

#### *Data Processing*

[A7] To make data processing more manageable for the purpose of the present analysis, the number of relevant points of motion was reduced by focusing on a single marker, namely the wrist joint. This choice is justified by the following observations:

1. Movements of the wrist can be isolated from the rest of the arm and other parts of the body; at the same time, when another body part is engaged—for example, the shoulder or the torso—the hand will move as well. It is thus a good indicator of both local and global movements. We thus ensured a continuous record of movement, even when that movement was not the result of engaging the hand directly.
2. The hand has relatively low mass and high energy, which means that one is capable of making a diverse range of movements at different temporal resolutions—from fast and small to large and slow gestures. As such, it was assumed that participants would engage it in all the musical contexts with which they were presented.
3. Hands have a privileged role in our everyday interactions with the environment, not only in manipulating objects around us, but also in communication, where manual gestures have been found to affect the very way we think (Goldin-Meadow 2003).

[A8] Movement analysis was based on the second recording for each excerpt, when such a recording was available. In cases where it was not, the first recording was used instead. Due to occasional marker occlusions, some recordings had to be excluded from analysis.

[A9] Movement analysis focused on acceleration as an indicator of movement co-occurring with sound. Absolute acceleration was calculated as a Euclidean distance between the successive derivatives (velocity) of the marker position data.

Raw data was processed in MATLAB using a smoothing algorithm and “gap-filling,” which is a method of filling small gaps in data by interpolating between the first and last missing frame using a piecewise cubic Hermite spline function with the preceding and succeeding frames as reference (Nymoer et al. 2012). Gap-filling was applied to gaps that were shorter than 20 frames (200 milliseconds) in length, discarding recordings with longer gaps. Altogether, this resulted in between 37 and 42 recordings of different participants in total per excerpt.

---

**Mariusz Kozak**  
**Columbia University**  
**Department of Music**  
**816B Dodge Hall, MC1813**  
**2960 Broadway**  
**New York, NY 10027**  
[m.kozak@columbia.edu](mailto:m.kozak@columbia.edu)

---

## Works Cited

- Abbate, Carolyn. 2004. “Music: Drastic or Gnostic?” *Critical Inquiry* 30 (3): 505–36.
- Berger, Harris. 2009. *Stance: Ideas about Emotion, Style, and Meaning in the Study of Expressive Culture*. Wesleyan University Press.
- Bispham, John. 2006. “Rhythm in Music: What is it? Who has it? And why?” *Music Perception* 24 (2): 125–34.
- Brown, Nicholas. 2006. “The Flux Between Sounding and Sound: Towards a Relational Understanding of Music and Embodied Action.” *Contemporary Music Review* 25 (1): 37–46.
- Cameron, Daniel J. and Jessica A. Grahn. 2014. “Neuroscientific Investigations of Musical Rhythm.” *Acoustics Australia* 24 (2): 111–16.
- Carman, Taylor. 1999. “The Body in Husserl and Merleau-Ponty.” *Philosophical Topics* 27 (2): 205–26.
- Castiello, Umberto. 2005. “The Neuroscience of Grasping.” *Nature Reviews Neuroscience* 6: 726–36.
- Certeau, Michel de. 1988. “Walking in the City.” In *The Practice of Everyday Life*, trans. Steven Randall, 91–110, University of California Press.
- Cimini, Amy. 2012. “Vibrating Colors and Silent Bodies: Music, Sound and Silence in Merleau-Ponty’s Critique of Dualism.” *Contemporary Music Review* 31 (5–6): 353–70.
- Clarke, Eric. 2005. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press.
- Cox, Arnie. 2011. “Embodying Music: Principles of the Mimetic Hypothesis.” *Music Theory Online* 17 (2).
- Cross, Ian. 2004. “Music and Meaning, Ambiguity, and Evolution.” In *Musical Communication*, ed. Dorothy Miell, Raymond MacDonald, and David J. Hargreaves, 24–44. Oxford University Press.
- . 2009. “The Evolutionary Nature of Musical Meaning.” *Musicae Scientiae* 13 (2): 179–200.
- Cusick, Suzanne G. 1994. “Feminist Theory, Music Theory, and the Mind/Body Problem.” *Perspectives of New Music* 32 (1): 8–27.
- Dreyfus, Hubert L. 2014. *Skillful Coping: Essays on the Phenomenology of Everyday Perception and Action*, with an introduction by Mark A. Wrathall. Oxford University Press.
- Fisher, George and Judy Lochhead. 2002. “Analyzing from the Body.” *Theory and Practice* 27: 37–67.
- Gallagher, Shaun. 2005. *How the Body Shapes the Mind*. Oxford University Press.
- Gibson, James J. 1977. “The Theory of Affordances.” In *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, ed. R. E. Shaw and J. Bransford, 67–82. Lawrence Erlbaum Associates.
- Godoy, Rolf Inge. 2006. “Gestural-Sonorous Objects: Embodied Extensions of Schaeffer’s Conceptual Apparatus.” *Organised Sound* 11 (2): 149–57.
- . 2010. “Gestural Affordances of Musical Sound.” In *Musical Gestures: Sound, Movement, and Meaning*, ed. Rolf Inge Godoy and Marc Leman, 103–25. Routledge.

- Goldin-Meadow, Susan. 2003. *Hearing Gesture: How Our Hands Help Us Think*. Harvard University Press.
- Goldin-Meadow, Susan, and David McNeill. 1999. "The Role of Gesture and Mimetic Representation in Making Language the Province of Speech." In *The Descent of Mind: Psychological Perspectives on Hominid Evolution*, ed. Michael C. Corballis and Stephen E. G. Lea, 155–72. Oxford University Press.
- Grahn, Jessica A. and Matthew Brett. 2007. "Rhythm and Beat Perception in Motor Areas of the Brain." *Journal of Cognitive Neuroscience* 19 (5): 893–906.
- Guck, Marion. 2006. "Analysis as Interpretation: Interaction, Intentionality, Invention." *Music Theory Spectrum* 28 (2): 191–210.
- Harrison, Daniel. 2011. "Three Short Essays on Neo-Riemannian Theory." In *The Oxford Handbook of Neo-Riemannian Theory*, ed. Edward Gollin and Alexander Rehding, 548–77. Oxford University Press.
- Hatten, Robert. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Indiana University Press.
- Hefling, Stephen. 1988. "Mahler's 'Todtenfeier' and the Problem of Program Music." *19th-Century Music* 12 (1): 27–53.
- Hirata, Catherine. 1996. "The Sounds of the Sounds Themselves: Analyzing the Early Music of Morton Feldman." *Perspectives of New Music* 34 (1): 6–27.
- Ingarden, Roman. 1986. *The Musical Work and the Problem of its Identity*. Translated by Adam Czerniawski. Edited by Jean G. Harrell. University of California Press.
- Jeannerod, Marc. 1988. *The Neural and Behavioral Organization of Goal-Directed Movement*. Oxford University Press.
- Jensenius, Alexander, et al. 2010. "Musical Gestures: Concepts and Methods in Research." In *Musical Gestures: Sound, Movement, and Meaning*, ed. Rolf Inge Godøy and Marc Leman, 12–35. Routledge.
- Kane, Brian. 2011. "Excavating Lewin's Phenomenology." *Music Theory Spectrum* 33 (1): 27–36.
- Kelly, Sean. 2002. "Merleau-Ponty on the Body." *Ratio (new series)* 15: 376–91.
- Koozin, Timothy. 2011. "[Guitar Voicing in Pop-Rock Music: A Performance-Based Analytical Approach](#)." *Music Theory Online* 17 (3).
- Kozak, Mariusz. 2012. "Moving in Time: The Role of Gesture in Understanding the Temporal Organization of Music." PhD diss., University of Chicago.
- Kirshner, Sebastian, and Michael Tomasello. 2009. "Join Drumming: Social Context Facilitates Synchronization in Preschool Children." *Journal of Experimental Child Psychology* 102: 299–314.
- Krueger, Joel. 2011. "Doing Things with Music." *Phenomenology and the Cognitive Sciences* 10 (1): 1–22.
- Leman, Marc. 2008. *Embodied Music Cognition and Mediation Technology*. MIT Press.
- Lewin, David. 2006. "Music Theory, Phenomenology, and Modes of Perception." In *Studies in Music with Text*, 53–108. Oxford University Press. Originally published in *Music Perception* 3 (4) (1986): 327–92.
- . 1987. *Generalized Musical Intervals and Transformations*. Yale University Press.
- . 1993. *Musical Form and Transformation: Four Analytical Essays*. Yale University Press.
- Lochhead, Judy. 2006. "Visualizing the Musical Object." In *Postphenomenology: A Critical Companion to Ihde*, ed. Evan Selinger, 67–86. SUNY Press.
- London, Justin. 2012. *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford University Press.
- Luck, Geoff, and Petri Toiviainen. 2006. "Ensemble Musicians' Synchronization with Conductors' Gestures: An Automated Feature-Extraction Analysis." *Music Perception* 24 (2): 189–200.
- MacRitchie, et al. 2013. "Inferring Musical Structure Through Bodily Gestures." *Musicae Scientiae* 17: 86–108.
- Marratto, Scott L. 2012. *The Intercorporeal Self: Merleau-Ponty on Subjectivity*. SUNY Press.
- McNeill, William. 1995. *Keeping Together in Time: Dance and Drill in Human History*. Harvard University Press.
- Mead, Andrew. 2002. "Bodily Hearing: Physiological Metaphors and Musical Understanding." *Journal of Music Theory* 43 (1): 1–19.

- Merleau-Ponty, Maurice. [1945] 2012. *Phenomenology of Perception*. Translated by Donald A. Landes. Routledge. Originally published as *Phénoménologie de la perception* (Paris: Éditions Gallimard, 1945).
- . 1964. “Eye and Mind.” In *The Primacy of Perception*. Translated by Carleton Dallery. Northwestern University Press.
- Metzer, David. 2009. *Musical Modernism at the Turn of the Twenty-First Century*. Cambridge University Press.
- Montague, Eugene. 2012. “[Instrumental Gesture in Chopin’s Étude in A-flat Major, Op. 25, No. 1.](#)” *Music Theory Online* 18 (4).
- Morris, Robert. 1995. “Compositional Spaces and Other Territories.” *Perspectives of New Music* 33 (1–2): 328–58.
- Moshaver, Maryam A. 2012. “Telos and Temporality: Phenomenology and the Experience of Time in Lewin’s Study of Perception.” *Journal of the American Musicological Society* 65 (1): 179–214.
- Muldoon, Mark. 2006. *Tricks of Time: Bergson, Merleau-Ponty, and Ricoeur in Search of Time, Self, and Meaning*. Duquesne University Press.
- Nattiez, Jean-Jacques. 1982. “Varèse’s ‘Density 21.5’: a Study in Semiological Analysis.” Translated by Anna Barry. *Music Analysis* 1 (3): 243–340.
- Noë, Alva. 2012. “What Would Disembodied Music Even Be?” In *Bodily Expression in Electronic Music: Perspectives on Reclaiming Performativity*, ed. Deniz Peters, Gerhard Eckel, and Andreas Dorschel, 53–60. Routledge.
- Noland, Carrie. 2009. *Agency and Embodiment: Performing Gestures/Producing Culture*. Harvard University Press.
- Nymoen, Kristian, et al. 2012. “A Statistical Approach to Analyzing Sound Tracings.” In *Post-proceedings of Computer Music Modeling and Retrieval*, ed. Solvi Ystad, 120–45. LNCS. Springer.
- . 2013. “Analyzing Correspondence Between Sound Objects and Body Motion.” *ACM Transactions on Applied Perception* 10 (2): Article 9.
- Phillips-Silver, Jessica and Laurel J. Trainor. 2007. “Hearing What the Body Feels: Auditory Encoding of Rhythmic Movement.” *Cognition* 105 (3): 533–46.
- Poudrier, Ève. 2012. “Multiple Temporalities: Speed, Beat Cues, and Beat Tracking in Carter’s Instrumental Music.” The Society for Music Theory/American Musicological Society/Society for Ethnomusicology Conference. New Orleans, LA.
- Reybrouck, Mark. 2001. “Biological Roots of Musical Epistemology: Functional Cycles, Umwelt, and Enactive Listening.” *Semiotica* 134: 599–633.
- Rings, Steven. 2011. *Tonality and Transformations*. New York: Oxford University Press.
- Roeder, John. 1994. “Voice Leading as Transformation.” In *Musical Transformation and Musical Intuition: Essays in Honor of David Lewin*, ed. Raphael Atlas and Michael Cherlin, 41–58. Ovenbird Press.
- Romdenh-Romluc, Komarine. 2010. *Merleau-Ponty and Phenomenology of Perception*. Routledge.
- Sanders, John. 1993. “Merleau-Ponty, Gibson, and the Materiality of Meaning.” *Man and World* 26: 287–302.
- Schettino, Luis F, Sergei V. Adamovich, and Howard Poizner. 2003. “Effects of Object Shape and Visual Feedback on Hand Configuration during Grasping.” *Experimental Brain Research* 151 (2): 158–66.
- Schiavio, Andrea. 2012. “Constituting the Musical Object.” *Teorema* 31 (3): 63–80.
- Schiavio, Andrea and Damiano Menin. 2013. “Embodied Music Cognition and Mediation Technology: A Critical Review.” *Psychology of Music* 41 (6): 804–14.
- Small, Christopher. 1998. *Musicking: The Meanings of Performing and Listening*. Wesleyan University Press.
- Smeets, Jeroen B.J. and Eli Brenner. 1999. “A New View on Grasping.” *Motor Cognition* 3: 237–71.
- Todes, Samuel. 2001. *Body and World*. MIT Press.
- Tomasello, Mark. 2008. *Origins of Human Communication*. MIT Press.
- Urista, Diane. 2003. “[Beyond Words: The Moving Body as a Tool for Musical Understanding](#).” *Music Theory Online* 9 (3).

- van Elk, Michiel, Marc Slors, and Harold Bekkering. 2010. "Embodied Language Comprehension Requires an Enactivist Paradigm of Cognition." *Frontiers in Psychology* 1, article 234.
- Voeglin, Salomé. 2010. *Listening to Noise and Silence: Towards a Philosophy of Sound Art*. Continuum.
- Volgsten, Ulrik. 2012. "The Roots of Music: Emotional Expression, Dialogue, and Affect Attunement in the Psychogenesis of Music." *Musicae Scientiae* 16 (2): 200–16.
- Wilson, Andrew D. and Sabrina Golonka. 2013. "Embodied Cognition Is Not What You Think It Is." *Frontiers in Psychology* 4, article 58.
- Wiskus, Jessica. 2013. *The Rhythm of Thought: Art, Literature and Music after Merleau-Ponty*. University of Chicago Press.
- Zbikowski, Lawrence. 2008. "Dance Topoi, Sonic Analogues and Musical Grammar: Communicating with Music in the Eighteenth Century." In *Communication in Eighteenth-Century Music*, ed. Danuta Mirka and Kofi Agawu, 283–309. Cambridge University Press.
- . 2011. "Musical Gesture and Musical Grammar." In *New Perspectives on Music and Gesture*, ed. Anthony Gritten and Elaine King, 83–98. Ashgate.
- . 2013. "Listening to Music." In *Speaking of Music: Addressing the Sonorous*, ed. Keith Chapin and Andrew H. Clark, 101–19. Fordham University Press.

---

## Footnotes

\* I'm grateful to Lawrence Zbikowski, Brian Kane, Jonathan DeSouza, Richard Hermann, and Marion Guck, as well as the anonymous reviewers of *Music Theory Online*, for providing insightful feedback and suggestions on earlier drafts of this article. I also wish to thank the editors of *MTO* for their help in getting this article to its final version.

[Return to text](#)

1. For commentary on Mahler's ambivalence toward programmatic music, see, among others, [Hefling 1988](#).

[Return to text](#)

2. Examples abound, some of which will be discussed below. Others include [Small 1998](#), [Clarke 2005](#), [Krueger 2011](#), and [Volgsten 2012](#).

[Return to text](#)

3. Statements about representations and their external manifestations cannot be validated empirically, so we need to take the authors' claim that their experiment "confirms" this observation with some skepticism. Furthermore, the authors do not consider whether the music, with the physical constraints it puts on the performer, is *causally* responsible for these pianists' gestures, rather than these gestures manifesting representations. Still, the point that there is *some* relationship between gestures and musical structure is well taken.

[Return to text](#)

4. Here, I use the term "listener" as a shorthand for participants in musical practices who are not themselves playing instruments or singing at the time of the performance. Not only is this a common locution, but I also find various alternatives (e.g., "audience member," or "non-performing participant," or "musicking non-performer") to be unwieldy and not entirely accurate in light of the different non-performing encounters with music that are possible ([Cross 2004](#)).

[Return to text](#)

5. For a review of different categories of gestures, including gestures of performers and non-performing participants, see [Jensenius et al. 2010](#). To the extent that such a separation is possible in practice, the authors suggest distinguishing between actions that are causally involved in producing sounds, and actions that are supportive of the former. For example, piano keys, which are immediately engaged in sound production, are typically depressed with the pianist's fingers, so those bodily actions are construed as "sound-producing" gestures. On the other hand, movements of the arms, torso, and the rest of the pianist's body participate in altering the intensity and articulation of sound, and in this taxonomy are considered "sound-facilitating."

[Return to text](#)

6. By highlighting the performative aspect of bringing sound into existence, Brown's approach echoes ideas put forth by early scholars of embodied music cognition, for example Godoy's (2006) *gestural-sonic objects*.

[Return to text](#)

7. Arnie Cox's "mimetic hypothesis" operates along the same lines, claiming that listeners understand music by overtly and/or covertly imitating movements of performers, and generating "motor mimetic images"—essentially that "thinking



about music involves imagining *doing* (making) music” (2011; emphasis original). Such a view not only places undue emphasis on imitation as a crucial mechanism in comprehension—a point that has been critiqued by van Elk et al. (2010), among others—but also obscures the nuanced way in which individual differences between listeners’ histories help shape musical understanding.

[Return to text](#)

8. For example, Alva Noë (2012) has argued that for him, as someone who does not play music, the phenomenology of musical embodiment does not seem to draw on performers’ gestures.

[Return to text](#)

9. See Wilson and Golonka (2013) for a general critique of what I here call the “quasi-embodied” perspective. Their main claim is that a fully embodied account of perception must consider the body as playing a “compulsory, critical, constitutive role” in solving a task at hand, and not simply as calibrating inputs and outputs from the brain. This latter view—exemplified by such approaches as conceptual metaphor theory, Fisher and Lochhead’s account, and Godøy’s statement above—claims that internally generated concepts are grounded in, and modified by, simulations of previous experience, which may or may not include embodied elements. In contrast, for a fully embodied listener the body must be a necessary and inevitable component of perception.

[Return to text](#)

10. Moreover, Lewin’s perspective on phenomenology is receiving increased attention in critical literature, making my attempt all the more timely. Some more relevant recent exegeses of Lewin’s [1986] 2006 article include Kane 2011 and Moshaver 2012.

[Return to text](#)

11. For a different take on musical spaces, see Morris 1995.

[Return to text](#)

12. For more on *qualia* in the context of musical materials, see Rings 2011.

[Return to text](#)

13. To give but one example, Roeder (1994) defines voice leading in terms of transformational technology. The upshot of his definition is that simultaneously attacked pitches can be construed as a single unified Gestalt, rather than a collection of discrete pitches, and voice leading can be modeled as a single operation rather than a collection of intervals.

[Return to text](#)

14. Daniel Harrison (2011) offers a different perspective, claiming that even if nothing literally moves in music, the metaphor of motion is nevertheless so evocative and widespread in theoretical discourse that it brings in a number of possibilities, contradictions, and points of interest to music analysis. Specifically, he directs our attention to the dichotomy between action and objecthood that this metaphor invokes. Such a dichotomy is predicated on a notion that musical objects in and of themselves are inert and have no tendencies. Only after transformations are applied as an “external force” can they participate in any kind of motion. However, they do so at the expense of a contradiction: the differences between types of objects (e.g., pitches, dynamics, articulations, chords, and so forth) are either disregarded, or else they are too rigidly determined. For Harrison, in contrast, musical objects are already imbued with tendencies, and so are themselves capable of performing actions.

[Return to text](#)

15. It will become clear in the course of this article that I think Lewin basically got it right, but was reluctant to make the jump to post-Husserlian phenomenology that would have allowed listeners’ perceptions to serve as a model of musical insideness.

[Return to text](#)

16. Indeed, Lewin’s list of “musical behaviors” is quite extensive, and includes various modes of sound production—from whistling and humming, through banging on diverse objects and blowing through pipes, to playing in formal and informal ensembles—of composition—such as writing pieces of music, improvising, transcribing—and of dancing ([1986] 2006, 96–7). Notably absent is *listening* as its own category.

[Return to text](#)

17. As evidenced by Kane’s exegesis, Lewin did not explicitly draw on Merleau-Ponty’s phenomenology of embodied perception in his own work. In fact, there is room for conjecture that the very concept of perception at the time of Lewin’s “Phenomenology” article (early-1980s) precluded basing it on bodily movements, because it was ostensibly entrenched in a computational, and therefore disembodied, paradigm of mental processing. At the same time, as my claims hopefully illustrate, his thinking was not far removed from the point I’m trying to make here: that bodily activity is already mental activity, and so can successfully underpin the kind of analytical engagement that he envisioned.

[Return to text](#)

18. Compare this, for example, with a definition of *gesture* offered by Robert Hatten as “any energetic shaping through time” (2004, 132).

[Return to text](#)

19. For a slightly different view of music intentionality that draws on studies of mirror neurons, see [Schiavio 2012](#).

[Return to text](#)

20. For reviews see [Cameron and Grahn 2014](#), [Grahn and Brett 2007](#), and [Phillips-Silver and Trainor 2007](#).

[Return to text](#)

21. There is no historically documented indication that Gibson and Merleau-Ponty were aware of each other’s work, even though their ideas appear to have reached maturity roughly at the same time. Still, there are interesting, seemingly complementary connections between their epistemologies, which are elucidated in [Sanders 1993](#).

[Return to text](#)

22. And also time; however, the issue of temporality is well beyond the scope of the present article. For a commentary on Merleau-Ponty’s phenomenology of time, see [Muldoon 2006](#), [Romdenh-Romluc 2010](#), and especially [Marratto 2012](#).

[Return to text](#)

23. Concerned as he was with music’s co-constitutive power in relation to the performer, it is likely that Merleau-Ponty himself would have endorsed this view. See [Cimini 2012](#) for a close reading of Merleau-Ponty’s perspective on music in regard to Cartesian mind–body dualism.

[Return to text](#)

24. For relevant evidence and discussion see [Cross 2009](#), [Kirshner and Tomasello 2009](#), and [Bispham 2006](#).

[Return to text](#)

25. An explication of this term can be found in [Wiskus 2013](#).

[Return to text](#)

26. Todes (2001), whose work greatly augments some of Merleau-Ponty’s key ideas, uses this as a basis for a downright virtuosic critique of Descartes.

[Return to text](#)

27. The two perspectives of experience and observations, and the concomitant categories of knowledge that they produce, are vividly juxtaposed in Michel de Certeau’s essay “Walking in the City” (1988), in which he compares the view of New York City from the top of the World Trade Center with the same at street level. One key difference, however, is that for de Certeau the players in the latter frame of reference produce an “urban text” that they themselves are unable to read, whereas in the view advanced here, and supported by Merleau-Ponty, insiders are generating a narrative that is completely intelligible to them. In other words, the “gnostic” viewpoint of the onlooker from up above is no more comprehensible than the “drastic” understanding of the participant down “in the trenches.”

[Return to text](#)

28. In addition to geometrical depth and depth in spatiality of situation, Merleau-Ponty also postulates “spectral depth,” which is a kind of depth evident in dreams and other liminal experiences of space ([Marratto 2012](#)).

[Return to text](#)

29. A similar point is made by Catherine Hirata (1996; see also discussion in [Guck 2006](#)) when she talks about sounds’ “integrity” resulting from their being “infused” with certain qualities by other sounds around them.

[Return to text](#)

30. Hubert Dreyfus has written extensively on “skillful coping,” which is precisely the kind of intelligent and purposive bodily activity in which we engage in our everyday dealings with the world. His numerous essays are collected in [Dreyfus 2014](#).

[Return to text](#)

31. This is borne out in experimental studies, for instance in [Schettino, Adamovich, and Poizner 2003](#). For more on how goal-directed movement is organized at the neural and behavioral levels, see [Jeannerod 1988](#). For more recent behavioral and neural models, see [Smeets and Brenner 1999](#) and [Castiello 2005](#), respectively.

[Return to text](#)

32. For reviews see [Godøy 2010](#) and [Nymoen et al. 2013](#).

[Return to text](#)

33. My view here is similar to that of Roman Ingarden (1986), who distinguishes between auditory objects that are identical to acoustical properties of sound, and auditory aspects (or Gestalts) that are constituted by experiences of listeners.

[Return to text](#)

34. See [Goldin-Meadow and McNeill 1999](#) for a discussion surrounding the fascinating question of why humans use speech, and not gestures, as the vehicle of language. Their assertion is that speech as a communicative modality has a weakness in that it cannot convey messages mimetically, a function for which manual gestures are excellently suited. In consequence, from an evolutionary perspective mimetic encoding was left to the hands, while the segmented and combinatorial form was taken over by speech.

[Return to text](#)

35. The sentiment is echoed by Lewin ([1987](#), 87 and [1993](#), 44). In all cases, Lewin's formulation has the potential to serve as a safety valve that allows one to claim a conceptual fracture between what analysis is meant to accomplish and what sorts of listening abilities one employs in perception, just in case there is friction between the two perspectives. But while such a move is legitimate and perhaps even imperative in the kind of formalisms that interest Lewin, theories and analyses that start from the *body* will necessarily need to make perception-statements. Because they are dealing with experienced phenomena, rather than abstract symbols, their aims diverge from those of a formal theory of intervals, simultaneously freeing the theorist from having to reduce behavior to logical operations, and constraining observational potency to avoid rampant subjectivism.

[Return to text](#)

36. For a superbly lucid discussion of how exactly objects are constituted in Merleau-Ponty's phenomenology—in a process that unfolds from phenomena, through things, and onto full-fledged objects of experience—see [Romdenh-Romluc 2010](#).

[Return to text](#)

37. While I won't pursue such a trajectory herein, one might further contend that the perceiver also constitutes him or herself as a *subject* in relation to the perceived as an object. For an exposition of this idea see [Voeglin 2010](#), especially ch. 1 (“Listening”).

[Return to text](#)

38. Inter-corporeality plays a critical role in Merleau-Ponty's conception of subjectivity. See [Marratto 2012](#) for more on this topic.

[Return to text](#)

39. With regard to the body playing a mediating role in this way, see [Leman 2008](#). His approach, which actually revives and reinforces the long-eradicated Cartesian split between the mind and the body, has been cogently critiqued by [Schiavio and Menin 2013](#).

[Return to text](#)

40. For the sake of simplicity only data collected from participants with musical training were used in the present analyses. See Appendix for additional information regarding methods and quantitative data analysis. For further commentary and discussion, see [Kozak 2012](#).

[Return to text](#)

41. See, for example, [Luck and Toiviainen 2006](#).

[Return to text](#)

42. On listeners' perception of multiple metrical streams in Carter's music see [Poudrier 2012](#).

[Return to text](#)

43. Participants were not aware of the titles and composers of the excerpts they heard.

[Return to text](#)

44. Because these movements, although regular, are not synchronized with the auditory signal, we cannot truly speak of *entrainment* in the sense of attunement to an underlying isochronous pulse ([London 2012](#)).

[Return to text](#)

45. One way to consider this is with Harris Berger's ([2009](#)) concept of *stance*: a mechanism of meaning formation that is not reducible to compositional techniques or to objectively determined features of music, but which comprises a complex position of the listener in relation to what they hear.

[Return to text](#)

46. By “figurative meaning” I am referring to the semiotic value of gestures, such as their communicative potential and subjective purpose.

[Return to text](#)

## Copyright Statement

Copyright © 2015 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online* (*MTO*) are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

Prepared by Cara Stroud, Editorial Assistant