1

Meaning

“Susanna” is the third movement of Andrew Norman’s string trio titled The Companion Guide to Rome (2010). This particular movement is a miniature for solo viola, and you can see it in its entirety in Example 1.1. One analytical observation that can be made right away is that it has a very sparse and rudimentary pitch structure, containing recognizable elements from common-practice tonality (such as chains of 4–3 suspensions and open fifths) without actually operating within a tonal system. It is somewhat reminiscent of J. S. Bach’s work’s for solo stringed instruments, but Norman does not use any identifiable quotations. In fact, whatever tonal techniques he does employ seem to be mere stock figures that could have come from just about anywhere, used more for their capacity to stand in as markers of archaism than for their motivic potential.

The whole-tone descent in the lowest voice supporting the aforementioned suspensions, and the unusual resolutions of the tritone—at the end of the first system (E–A♯ → E–B) and later in the middle of the second system (A♭–D → A♭–E♭)—together provide a way into the pitch and harmonic structure of the piece. However, an analysis that focused on only these elements would tell but a small part of the story, and would need to be supplemented by an account of how the piece’s meaning is partly constituted by the performer’s body. The violist is instructed to apply heavy pressure to the bow while initially shaking it and, later, moving it very slowly, producing sounds that barely escape the instrument. From an almost inaudible G♯–B dyad in the opening, to the full-throated broken chords in the third line of the score, there is a gradual opening of sound, an increase in clarity that corresponds with the upsurge of dynamics. A dominant-like C–B suspension against an

1 A recording with the score is available on YouTube at https://www.youtube.com/watch?v=40ZPD5b_tMDU (the relevant movement starts at 2:40).


III. SUSANNA

Start with the bow on the string... 

Before shaking, immunnitously at first... 

Let a few short sounds escape from the strings... 

(gently let more sounds escape until...) 

the bow is starting (as quietly as possible) back to shaking... 

...starting... 

Stop the bow at the frog... 

Shake... 

Gradually phase the shaking into shaking and the shaking into stillness... 

Shake. 

Place the bow (near the frog) on the strings with heavy pressure. 

Shake the bow in small and quick up and down motions while maintaining the same point of contact with the strings. 

One should be able to shake the bow and strings in this manner without letting any sound escape the instrument. 

Crescendos and diminuendos affect the amplitude of the shaking, and can be added at will. 

Gradually, as marked, let small short sounds escape as the shaking transforms to stammers. 

Shatter... 

Move the bow very slowly, with sufficient pressure such that the tone breaks up into little, irregular starts and stops. 

With the right amount of continuous pressure, these little sounds do not need separate impulses from the arm. 

Rather, they will emerge naturally from the “stammers” of the bow on the string. 

The speed of the stammer need not, and should not, be totally constant. 

The changing distances between vertical marks are used only to convey an overall effect of irregularity, not to convey specific speed changes in the stammer. 

Change bows freely as needed to maintain the effect.
open G–D fifth suggests imminent tonal closure, but the sound is arrested once again. Finally, in a last-ditch effort, the music lunges into an exasperated climax on a broken d-minor triad, only to be brutally and summarily choked by the violist’s heavy bow.

The body of the performer becomes implicated in the constitution of what I would like to call emergent and transient musical objects. These are objects that lack the kind of stability that characterizes objects that usually concern music analysts—pitches, chords, rhythms, and so on. Instead, they are fleeting phenomena that occur in the corporeal relationships between performers and listeners, arising and dissolving together with the upsurge of time. If they can be said to exist at all, they do so only in the sense that forces do: exerting influence without betraying their materiality. Yet despite their ephemerality and evanescence, they are nevertheless empirically real insofar as they are the felt qualities of one’s encounter with a piece of music.

In the case of “Susanna,” the body of the violist does not merely reproduce notated pitches, but quite forcefully conceals “normal” sounds behind the harshness and awkwardness of the stutters and the shakes. While it is true that these stutters and shakes can be represented symbolically—as the composer did in the score—their auditory effect evades visual apprehension, lingering only as a particular aspect of the intercorporeal network established during a performance. Rather than something extra-musical, the violist’s body is very much an indissoluble element of the music. Sounds function as vehicles for a body caught up in an action that urges musical interpretation. In light of this, a more thorough analysis might weave, for example, stories of how sounds signify particular kinds of bodily exertions, and what those exertions might, in turn, signify of the person producing them; or perhaps stories about St. Susanna, the third-century Christian martyr and patron of the Roman church who inspired this movement; or maybe even stories that critique and challenge our societal assumptions regarding bodily norms and abilities. In all these cases, there is a human agency latent within sound, a gesture that gives it life and becomes the (transient and emergent) object of analytical attention.

Let us delve a bit into this agency, this gesture: Who makes it? What is its musical significance? How do we incorporate it into our analytical stories?

The human body seems to stand in for someone who is more than a mere collection of his or her parts. A “self” or a “subject” appears to inhabit the body and give it character—playful, lustful, sick, angry, fragile, powerful, confident, timid, and so on. Below I will give some nuance to this claim, but for now let us continue with the ostensibly incontrovertible assumption that people have bodies, and what those bodies do is a direct expression of a concealed entity—call it the subject, the mind, consciousness—that guides them. In ordinary circumstances, subjectivity is on full display, manifested in the body’s postures and actions.

However, as Naomi Cumming (2000) has argued, musicians’ bodies signify this subjectivity somewhat differently, because they exist in a liminal space between pure physicality (exemplified by the sheer athleticism of technical facility) and pure musicality (mediated and interpreted as particular kinds of sonic signs). As such, it becomes possible to conflate sounds with musicians’ identities because “the characteristics of sound are the aural ‘marks’ of bodily actions” (22). Even without the visual cues of a live performance, “the impression of ‘personality’ can be gained subliminally through the markers in sound of what seem to be the performer’s characteristic physical responses” (22). On this view, when listening to a recording we do not merely attend to the sounds, but simultaneously construct a body that produced them (or one that we imagine could have produced them).

For Cumming this attention to something that exceeds the acoustical signal constitutes the heart of musical experience. For example, a so-called singing violin sound (as highly desirable as it is elusive) does not emerge because this sound merely refers to vocality, or because the violinist imitates the singing voice with the instrument. Rather, Cumming claims that listeners interpret it as emanating directly from the violin because “singing” is “heard as belonging to a sound” (2000, 75).

The pedagogical tradition of comparing sounds of stringed instruments to singing goes back at least to Leopold Mozart’s *Treatise on the Fundamental Principles of Violin Playing*, where he writes that “singing is at all times the aim of every instrumentalist” (Mozart [1756] 1951, 102). In the same

---

4 Although I do not explore this issue further, it is worth mentioning that this claim puts a wrinkle in the ecological approach to listening, according to which sounds encode their own means of production (see, for example, Clarke 2011). The question is where the quality of singing comes from, if the sound of the violin itself indexes its source: a bow scraped across a string stretched over a wooden box. It seems that the most fruitful way of approaching this problem would be through the lens of analogy, whereby there are certain characteristics of both vocal and violin sounds that map onto each other (see Zbikowski 2017).
passage—particularly relevant to Norman’s “Susanna”—he further praises the human voice for its ability to “[glide] quite easily from one note to another,” without creating a break between notes except to produce “some special kind of expression, or the divisions or rests of the phrase demand one” (102). Because of this long history of associations, the sounds that the violist makes in “Susanna” can be heard as violating a well-established norm, marking them as pathological (stuttering, choking, etc.). In turn, this allows the listener to create an image of a body that might produce these sounds, a body that struggles to express itself, a body engaged in some excruciatingly difficult and painful labor, fighting to liberate itself from whatever internal or external power is trying to suppress it.

But remember that for Cumming vocal pathologies belong to the sound and not to the body of the performer. This means that we are not dealing with the real, physical human body directly engaged in making sounds. The violist in “Susanna” is not literally choking or stuttering. Instead, Cumming proposes that performers project what she calls “presence,” which is a body created metaphorically through acts of interpretation. Bodily presence in sound is mediated by language or other representations—it is a sign. Or, to use another of Cumming’s terms, the sound conveys a “virtual agency,” akin to Edward T. Cone’s “persona” or Carolyn Abbate’s “figural subject” (Cone 1974; Abbate 1991). These agents, personae, and subjects all take the human body as their (imagined) form, but it is not a body made of flesh and bones. Instead, it is a body created in the semiotic act of listening, a body that is unrestrained by physical laws and hence capable of superhuman feats. In short, it is a body that has been defleshed and deboned. This somewhat grotesque act of butchery displaces the immediacy of communication between performers and listeners, turning it into an “illusion,” a “mediating representation” created by the performer’s negotiation of “the mediating space between physicality and interpreted gestural motion” (Cumming 2000, 160). Presence here is a construct, an effect of semiotic play.

Perhaps sometimes this semiotic play is necessary to create a distance between ourselves and the music, an act reminiscent of Homer’s Odysseus tying himself to the mast of his ship in order to experience the treacherous song of

---

5 Mozart’s larger point conforms to the prevalent naturalism of the Enlightenment, where the human voice was considered the most natural instrument available. While not writing specifically about singing, in his 1751 treatise, *The Art of Playing on the Violin*, the violinist and composer Francesco Geminiani claimed that “The Art of playing the Violin consists in giving that Instrument a Tone that shall in a Manner rival the most perfect human Voice” (quoted in Weiss and Taruskin 2008, 221).
the Sirens. (I will return to this image.) But contrary to Cumming, I propose that the presence of the body in “Susanna” is not created at a distance, over a culturally mediating space full of promises of safety and aesthetic enjoyment. We are not dealing here with a body that is a product of our imagination, nor is the intimacy enacted by the performer a mere effect of interpretive work, an extra-musical appendage in excess of the notes themselves. No, these are not the sounds of nobody: these are sounds made by real flesh and blood and stained by pathology and violence. The relation between performer and listener is immediate precisely because sounds mark the bodies that produce them: sounds and bodies are indissoluble.

Certainly, the nature of this relation may not be fully captured by metaphorical descriptions like “stuttering” and “choking.” When I hear this piece I do not have the urge to leap onto the stage and start administering the Heimlich maneuver to the (likely bewildered) violist. Nevertheless, I want to suggest that the whole point of “Susanna” is to hear the body that makes the sound, not just the one imagined in it; to hear the violist defy and defile those very modes of sound production that constitute our Western performance tradition; to hear her body tense up, fold up, force itself into shapes and gestures that transgress everything she has painstakingly cultivated through years of study. The communion thus established between the performer and her listeners is far from an illusion created by semiotic play, but instead is as real and as moving as those between bodies engaged in intimate—if violent—nonmusical acts. For, in a manner that is reminiscent of Luciano Berio’s Visage (1961 for voice and tape), there is something violent—almost brutal—in how sounds are wrought out of the instrument.

Like the performer of “Susanna,” I too play a string instrument, hence the brutality of the deed is likely to be more vivid for me than for a pianist, or a singer, or a listener who has never engaged in musical performance. At any rate, the intensity of the encounter is not at stake here. I raise it to illustrate how the physical communion between performer and listener brings forth those emergent and transient musical objects I mentioned in the beginning, objects that may not necessarily be concrete and precise, but are nonetheless experientially genuine and transformative. It does so because, in addition to the aforementioned commonly held view that a subject inhabits the body and is responsible for its doings, we also are our bodies. We do not merely have them; bodiliness is not just a matter of possession, as if the fleshy mass simply adhered to our disembodied “selves.” Instead, one of the central claims of this book—to be given fuller treatment in Chapter 3—is that our
subjectivity is constituted by the very bodies that we are, and that this has important consequences for how we consider time in music.

An approach that regards the intimate link between performer and listener as a mediating illusion—one that listens to the piece, as it were, from the comfortable distance of semiosis—would not create a sufficiently rewarding listening experience. In contrast to Cumming’s claims, the manner in which Norman directs the violist to perform the piece does not merely inscribe the body in the sound, but also makes it so that the sound is heard as explicitly issuing from a very real, physically present body. And it is the tangible corporeality—corporeal reality—of this body that becomes musically meaningful. The question, then, is how to harness this corporeal meaning and make it a part of our analytical stories.

**Musical Objects**

Part of the problem with including corporeal relations in our analyses is that these relations are not easily extricated from the time in which they unfold. Indeed, our ability to talk about musical processes is largely founded on our capacity to turn them into quasi-objects, which is why so much effort is put into finding convincing ways of segmenting music. The incessant flow of events has to be partitioned into chunks that are significant in and of themselves, outside of temporal passage. To be convincing and intersubjectively agreed upon, this partitioning must be rendered on the basis of justifiable criteria, such as changes in pitch, harmony, rhythm, contour, articulations, and so on (Hanninen 2012). Rather than having to rely on our fleeting and imperfect memories, we give coherence to our thoughts by removing music from flux. Only when music becomes a fixed entity, resistant to the disintegration that is characteristic of temporal objects, can we give it meaning by pointing to other, non-musical entities (Carpenter 1967).

This process of fixing and pointing requires that the entity in question endure on its own, or be made to endure in a way that facilitates transmission.

---

6 An important and influential example of this is Hanninen (2012). See also Hasty (1981).
7 Matthew Butterfield (2002) draws on the work in image schemata and categorization to ground Carpenter’s ideas in early twenty-first-century models of cognition. His claim is that “we tend naturally and spontaneously to map our experience of physical objects onto our experience of sounds and thereby ‘objectify’ them” (333). Having undergone this process, musical sounds can then guide listeners’ behavior. Here I offer a different perspective, wherein I draw on Gibsonian ecological perception to argue that behavior takes place prior to objectification.
Why this is the case can be gleaned from Cumming’s use of Peircean semiotics as a gateway into the formation of musical meaning. On this approach, sounds function as signs by pointing to things outside of themselves, but in order to be apprehended as such by a listener, the external reference has to be reproducible (Cumming 2000, 28). It must be somehow imprinted on the listener through repeated patterns of activity. It is this process of imprinting—this reproduction—that I appeal to when I say that we give coherence to our thoughts: a pattern of activity that carries a particular meaning has to have features that are recognizable from one iteration to another. For this to happen, the relevant pattern has to be temporally bounded, otherwise it would not be reproducible but would merely constitute a continuous production of something new—an idea I will take up in Chapter 6. Sound considered on its own has no such bounds; instead, these bounds have to be created by the listener, making segmentation a necessary aspect of music analysis. This is what I mean by enduring musical objects. Their creation is a process that has obvious advantages in various forms of communication. I call this a “referential” view, whereby “anything acquires meaning if it is connected with, or indicates, or refers to, something beyond itself, so that its full nature points to and is revealed in that connection” (Cohen 1946, 47). As shown in Figure 1.1, meaning here emerges from a process of triangulation: (1) the acoustical signal is a phenomenon to be interpreted by pointing to (2) a linguistic, symbolic, or gestural musical object for which it stands, created by (3) the interpretive act of a

\[\text{Figure 1.1 Triangulation of musical meaning.}\]

\[\text{\footnotesize \textsuperscript{8} See also Meyer (1957).}\]
listener who furnishes the conceptual framework for this particular relation by drawing on his or her personal and socio-cultural background.\footnote{A similar triangular model is found in Meyer (1956, 34).}

Referential approaches essentially create non-temporal musical objects that can be compared to other objects, stored for later retrieval, or manipulated using various operations. Here, I do not mean to suggest that the enduring object itself cannot have temporal features. For example, a musical object recognizable as an “authentic cadence” consists of a temporally ordered sequence of harmonies, and this temporal ordering is crucial for us to be able to identify the object as such. Rather, what I wish to emphasize is that this particular sequence acquires a kind of stability that allows it to endure on its own as a repeatable pattern. Its features can be identified with enough precision to accurately reproduce the sequence regardless of how it unfolds in real time. This reproducibility proves to be quite efficient when dealing with ephemera like music, because it facilitates the emergence of a joint attentional framework which can serve as a discursive base (Tomasello 1999). There is a stable relationship between sounds and extra-sonic representations, a relationship that remains long after the process has ended. Such referentiality works well whenever one finds substantial inter-opus agreement about the properties that the objects in question should have. For example, if we were interested in how a sonority that is identified as a “g-minor root-position triad” functions in different contexts, we need not concern ourselves with all of its characteristics, but only those that it has in common with other sonorities that can also be labeled as “g-minor root-position triads.” From this perspective, the opening chord of Bach’s Partita No. 1 for Solo Violin (BWV 1001) is of the same type as the chord found on beat three of m. 2 in the same piece, or even on the downbeat of m. 3 of Mozart’s Symphony No. 40, K. 550 (see Example 1.2). Our recognition of this relationship is grounded in our broader capacity for categorization, a basic cognitive skill that allows us to function efficiently in our environment (Zbikowski 2002). Rather than having to redefine each g-minor root-position triad from scratch, the process allows us to set aside what happens while the sound unfolds, and to focus our attention on other pertinent aspects, such as when a similar sonority occurs in the piece, or how much time it takes before it is supplanted by a different sonority.

Yet composers increasingly create sounds that are difficult to inscribe using existing representations, or to describe using commonly found terms. As a case in point, consider “60 Pieces of Sound” (2009) by Jürg Frey, an Austrian composer who is one of the members of the Cage-inspired Wandelweiser group.\footnote{A YouTube video of the piece can be found at https://www.youtube.com/watch?v=hVpPEf5dOmc.} Here, the
The listener is presented with sixty “blocks” of sound, each standing like a monolith separated from the others by long stretches of silence. If we focus on the plural noun in the title, then we might be led to believe that these sonorities are meant to be appreciated as somewhat autonomous units, perfectly suited to carry the designation of a “musical object.” The underlying assumption is that we are dealing with a single sound splintered into sixty parts. Yet the complexity of these sonorities, and, importantly, their changing quality over the course of their duration, injects some ambiguity into this interpretation. On the one hand, it is possible to think of the pieces as arranged linearly, like dominos, such that putting them together would result in one long, metamorphosing sound. This hearing would focus on the connection between successive sounds, filling each silence with anticipation of how the just-heard process might continue. On the other hand, it is equally plausible to consider each “block” as part of a three-dimensional object that has been broken up into fragments, “originally” having sounded together. Here, instead of linear development, our attention would be on music as it folds.

**Example 1.2a**  J. S. Bach, Sonata for Solo Violin in G-minor, BWV 1001, Adagio, mm. 1–2.

**Example 1.2b**  W. A. Mozart, Symphony No. 40 in G-minor, K. 550, mm. 1–7.
over itself like an origami sculpture, requiring us to remember—or at the very least have a feel for—how the pieces fit together.

This interpretive ambiguity results from the fact that the unfolding of these sonorities is coexistent with the time it takes us to experience them; they are never present in their entirety. Only when their sound disappears can anything approaching a “musical object” emerge through the referential process of semiotic triangulation—but by then our attention is already drawn to something else, making it impossible to establish a definitive relationship between the titular pieces. Like the emergent and transient intercorporeal relations in “Susanna,” the sounds of “60 Pieces” cannot be extricated from their becoming without losing something of their essence.\footnote{The way I conceive of the inextricability of time from the sounding object is similar to Robin James’s (2010) “conjectural body,” which is her term for the notion that bodies are “inseparably material and social,” hence any attempt to separate out different categories of social identity ends up dealing with phenomena that “never actually exist as such” (xiv).} Useful in this regard is Helmut Lachenmann’s distinction between sounds that can be perceived as objects and sounds that can be perceived as processes. In the former case, a sound object achieves its full characteristic expression before the auditory signal ends—think, for example, of a static sustained chord. The object’s features are exhausted while it is still sounding, and once they have been identified, we no longer have to attend to this object’s unfolding. By contrast, a sound process lasts precisely as long as its duration in real time; it unfolds without fully revealing itself (Lachenmann 1996; in Tsao 2014). A process must be considered as a whole in tandem with its duration. Put differently, attending to a process requires a consideration of its temporal nature. Frey’s sonic blocks, which are characterized by continual changes of timbre and intensity, belong in this latter category.

Although the main focus of this book is more recent repertory—a choice driven largely by opportunities for developing novel analytical approaches—the notion that processes must be regarded as inextricably linked with their temporal unfolding also extends to the music of other historical periods, and includes the Bach and Mozart examples mentioned earlier. To consider the former, in addition to identifying the opening chord of BWV 1001 as a g-minor root-position triad, one would be remiss not to attend to the special technique that a violinist has to employ in order to “break” the chord. Due to the physical structure of the instrument, the sonority must gradually morph from an open-string fifth (G–D), which rings unencumbered as it supports the addition of the mode-defining third (B♭), all of which then disappear, one
by one, until what is left sounding is the high G5.¹² The process can be languorous and meditative (as in Anna Göckel’s contemplative rendition¹³), or brisk and incisive (an interpretation favored by Shlomo Mintz¹⁴ and Nathan Milstein¹⁵), but in each case the highest pitch emerges as if from a lifting fog, achieved through a process, rather than freely given. It is the very first sound we hear in this piece, which sets up a contrast with the similarly labeled chord in the third measure of K. 550. There, the “g-minor root-position triad” is a repetition of what came earlier, a reiteration in the midst of the ongoing stream that eventually coalesces into the symphony’s primary theme. Unlike the Bach sonority, it is a true simultaneity: the pitches that give this chord its name are present all at once. And yet it too is not an object. Like the temporally spread out opening of BWV 1001, it is a process, unfolding with characteristic breathlessness, its urgency underscored by the agitated arpeggiation in the violas and the repeated upper-neighbor motif in the violins. Indeed, this triad is constitutive of an even larger process, one which establishes the g-minor as the home key and only changes with the arrival of a new harmony—a half-diminished seventh-chord in third inversion—on the downbeat of m. 5.

The point is that these intangible musical processes resist being crystallized in symbols. Over the centuries, of course, composers, performers, and theorists have developed ways of doing precisely that, and one could easily come up with notation that would help with categorizing the processes heard in “60 Pieces” and “Susanna.” Resistance can be overcome—by force if necessary—but such overcoming voids those processes of any on-the-edge indeterminacy, making them listless and barren. I am not referring here to the symbolic representation that is the purview of composers—scores in the Western musical tradition serve to impart instructions to performers, and are thus an essential element of how music is conceived. Rather, what concerns me is the analytical method itself—the way in which the auditory signal is recaptured in order for the analyst to come to a fuller understanding of its meaning. To identify musical processes by their function, or to assign

¹² The relationship between instrument design and aspects of musical structure is discussed in De Souza (2017). I will return to it in subsequent chapters.
to them unique labels, is to impose on them an ontological stability that inevitably forestalls the flow of time. Assigning an identity to musical processes renders them impervious to flux: having been given a function, they endure even after their material existence as sound pressure waves has long faded into the past. Although such assignation allows us to treat sound processes like physical objects—manipulating and transporting them across domains in ways similar to chords, pitches, harmonic functions, and intervals—doing so asserts a value that inevitably and irrecoverably changes their nature by attempting to mitigate (instead of embracing) their temporality. Time can subsequently be re-injected into the analytical model, but by then it has lost much of its efficacy. The problem is one of conceptualizing both time and music in a way that keeps them interwoven.

**Objective Time**

To do that—to embrace the temporality of musical processes—is to heed Christopher Hasty’s admonition to take time seriously (Hasty 1997, 303). It is to think of the ontological status of musical processes as fundamentally constituted by how they become, how the very ongoingness of time and its refusal to remain fixed inheres in them as an inextricable aspect of their being.\(^\text{16}\) Yet this task is not so easily accomplished. Despite Jonathan Kramer’s claim that music “becomes meaningful in and through time” (1988, 1), Victor Zuckerkandl’s affirmation that it is “the temporal art par excellence” (1956, 151; see also Hasty 1997, 20), or Robert Morgan’s even more emphatic declaration that “there is no question, of course, that music is a temporal art” (1980, 527), Norman’s and Frey’s pieces offer two different illustrations of how the relationship between musical meaning and time can be equivocal.\(^\text{17}\) In the former case, the ephemerality and evanescence of the network of relationships between a performer and her listeners makes it difficult to extricate the resulting corporeal sensibility from temporal flow without jeopardizing the nuanced ways in which real, fleshy bodies constitute musical meaning. Time in “Susanna” is not just a container within which the piece unfolds; it is created by the bodily interactions of participants. Meanwhile,

\(^{16}\) Further clarification of this idea can be found in Hasty (1999).

\(^{17}\) Claims like these are by no means the product of the most recent history. Jean-Jacques Rousseau ([1781] 2000) already remarked in his *Essay on the Origins of Language* that “the field of music is time.” For other eighteenth- and nineteenth-century antecedents of the assertion that music is primarily and thoroughly temporal, see Taylor (2016).
“60 Pieces” is made up of blocks that reveal themselves to be in flux, making time their constitutive, inextricable dimension. Once again, the form of sound emerges in the process of its unfolding, not as a temporal shape that exists beforehand.

Perhaps there is a fair degree of contingency at work in proclamations of music’s temporality, insofar as the tools for analyzing musical time have been largely developed in the context of common-practice tonality. There are certain long-standing assumptions—drawn most notably from the work of Newton and Descartes, whose ideas seeped into music-theoretical discourse in the eighteenth century—that analysts tend to make about the nature of time, and about the concomitant compositional techniques that bring this nature to light.\(^{18}\) Consisting of both harmonic and rhythmic elements, these techniques have come to use certain progressions or metrical patterns to represent specific features of time. For example, harmonic progressions that lead to cadences can be said to signify temporal linearity, while progressions that meander without a clear goal can indicate timelessness; repetition of earlier material can intimate a return to some prior moment in time; sequences have the capacity to suspend the forward momentum of time; and rhythmic augmentation makes the time go slower, while diminution speeds it up.

This contingency of temporality, as I mentioned with regard to Bach and Mozart, says nothing about the musical practice itself—which is as vitally temporal as any non-tonal work from the last hundred years—but it does highlight the problematic foundation of the concepts that analysts use to talk about it. Underlying these concepts is a perspective that posits two independent, well-defined entities: time and music. Music is said to unfold in time—a statement believed to be so unassailable that, more often than not, it is simply asserted as fact.\(^{19}\) Music here depends on time for its meaning, because it needs temporality as its vector. In turn, it has the capacity to alter temporal experience because this experience is predicated on a stable link between musical and temporal concepts. This view is very much in keeping with the common-sense understanding of time as something that exists regardless of its contents, as evidenced by Thomas Mann’s beautiful description: “Into a section of mortal time music pours itself, thereby inexpressibly enhancing and ennobling what it fills” (1924). Ever since Galileo emancipated time from movement, it became

\(^{18}\) For Newton’s influence on eighteenth-century music theorists’ conceptions of time—especially as regards meter—see Grant (2014) and Hasty (1997).

\(^{19}\) One need only do a quick online search to see just how many sources state this claim and its cognate, that “music exists in time.”
the fourth dimension—an independent variable on par with space. Newton, with his notion of “absolute, true, mathematical time,” further endowed it with a unity and unalterability that thoroughly separated it from all other universal forms, particularly motion and change (Newton 1972). By linking geometrical points with a line, Western ideas and representations of time voided it of any specific matter, and hence of a need for a witness to its passage. Rather than singling out a unique moment as the present, with the past trailing behind and the future encroaching from the opposite side, absolute and objective time found itself without even so much as a “now,” supplanting it with an anonymous, unconditional label like a notch on a measuring stick: at $t = x, y$, where $t$ is time, $x$ its measure (numerical value), and $y$ is some event.

The specific topology of Newton’s time has since been challenged from within the field of physics, most famously by Einstein’s theory of Special Relativity and later by quantum mechanics (Rindler 2001). Yet despite these deformations, what has remained intact is the notion that time exists independently of events that occur in it. This kind of time, which for physicists serves as a measurement of the rate of change of position, goes by different names: objective, physical, public. All of these labels refer to a time that is broken up into segments of nested sizes, from nano- and micro-seconds, up to geological periods, epochs, eons, and beyond to supereons. It includes purely theoretical time chunks, like Planck time, which is about $5.39 \times 10^{-44}$ seconds, or the time it takes light to travel 1 Planck length (Faber 1987). At every stage, from the infinitesimally tiny to the immeasurably huge, objective time is markedly atomized, precisely quantifiable, and obstinately indifferent to just about everything else.

There are, of course, undeniable benefits to this conception of time. Perhaps the most significant is our ability to precisely model phenomena outside of those directly related to our experience—such as the shape and makeup of far-off galaxies, or the conditions on Earth prior to the appearance of life—to say nothing of Planck time and space. Meanwhile, closer to our everyday concerns, the development of a public frame of temporal reference that can be used to coordinate and regulate the behavior of a large number of individuals has had an immeasurable impact on the socio-economic landscape of the Western world (Landes 1983). Much of the advantage comes from the spatial qualities which can be projected onto time, and from the resultant

---

20 For a thorough and lucid review of the historical developments in the Western conception of time, from its origin in Aristotle’s Physics, through Newton and Einstein, up to Heisenberg, see Lochhead (1982). Taylor (2016) provides a digest of the philosophical debates, particularly in the eighteenth and nineteenth centuries, and their reflection in musical discourse.
Meaning

isomorphisms between temporal chunks and the mathematical properties of serial continuity (Whitehead et al. 1919). We speak of moments in time and movement through time; events either come toward us or turn into places toward and away from which we move. We use statements like “The time will come when . . . ,” or “We are coming up on Christmas,” both of which make use of metaphors that conceptualize time in terms of motion (Lakoff and Johnson 1980). Time becomes an entity with a physical extension that can be conceived as a bounded, measurable region. We treat instants as if they were containers for completed wholes—“music is meaningful in time.” According to Mark Muldoon, this conception is indebted to the Enlightenment’s fascination—obsession, even—with the scientific method, and the desire to understand nature as “a unity held together by unassailable laws” (2006). Vagaries of subjectivity have no place in an episteme that not only seeks to identify various generalizations underlying the particulars of everyday life, but also attempts to cleanse them of what Frances Dyson (2009) calls “ontological orphans”—immateriality, invisibility, and ephemerality.

The benefits of having an objective notion of time extend into music theory and analysis, enabling a variety of observations and assertions about musical processes that otherwise remain beyond our grasp. To take but one example, the professed isomorphism between objective time and the serial order of numbers, in combination with the cyclical property of musical meter, has allowed theorists to conceptualize rhythm in terms of beat classes and beat-class sets, and to employ in music analysis operations familiar from twelve-tone theory, such as transposition and inversion. On this view, beat-class sets are formally analogous to pitch-class sets—which themselves generalize collections of pitch-classes—and denote which beat in a metrical pattern is articulated. As described by Richard Cohn (1992), “like pitch-class sets, [beat-class] sets have an interval content, bear properties such as invariance or cycle-generability, and enter into equivalence, similarity, and inclusion relations with each other. Consequently, much of the technology developed for

21 Whitehead had a tremendous distrust of the ability of language to explain temporal phenomena in any way that was close to experience. See more in Mays (1972).

22 Arnie Cox (2017) analyzes the use of temporal metaphors in music theory and analysis.

23 On music theoretical applications of Foucault’s epistemes, see Moreno (2004).

24 The isomorphism is actually a bit more complex, because conceptions of rhythm in terms of arithmetic sets draw on a more general analogy between rhythm and pitch, which has already been connected with numerical properties (Babbitt 1962). For analytical applications see especially Cohn (1992) and Roeder (2003).
Enacting Musical Time

atonal pitch-class analysis is transferable to the rhythmic domain, *mutatis mutandis*” (149). This method of conceptualizing meter in turn illuminates properties of temporal patterns that are not readily available to perception, such as formal similarities between seemingly disparate rhythms.

However, the concept of beat classes essentially renders musical rhythm a-temporal by allowing it to be taken out of the changing musical context and assigned to some category on the basis of its distribution within a measure, as if the latter were a kind of abstract, endlessly reproducible container. This is a byproduct of a more general tendency to regard meter as a grid into which the musical process is “poured” (to reiterate Mann’s poetic image). As noted by Roger Grant (2014) in his discussion of the temporal revolution that swept European music theory in the wake of Newton’s development of calculus, the eighteenth century saw a “shattering” of the theoretical “edifice that had once joined meter, character, and tempo” (93). In particular, pre-Newtonian Scholastic philosophy and science envisioned time as fundamentally intertwined with motion. In music theory, which belonged part and parcel to this scientific tradition, the connection was evident in the embodiment of the beat—an explicitly physical act used to “measure” music. As Grant argues, the act of “beating time”—the up-and-down motion of the hand—was critical in solving an ancient temporal dilemma: to think of time as simultaneously continuous and consisting of separate instants. By contrast, the new science developed by Newton, and its transmission into music theory via Johann Phillip Kirnberger’s treatise *Die Kunst des reinen Satzes in der Musik* (1771), divorced musical meter from its physical embodiment, grounding it first in the notated measure and then in mental accentuation of an ongoing series of otherwise undifferentiated beats, before finally becoming an aesthetic activity of organizing musical flow (Grant 2014).

Throughout these transformations, time remained independent of physical motion, constituted by its own succession of autonomous instants.

The terms of the above discussion are deeply entrenched in modern-day thought, even if on some level “objective time” can be recognized as just a theoretical construct. In his penetrating phenomenological account of the experience of music, Thomas Clifton (1983) was categorical in his refutation of the reality of objective time: “Objective time (or real, or absolute time) is a contradiction in terms. It presupposes the existence of a time which exists

---

25 This latter view has recently enjoyed a resurgence, most notably in the work of Justin London (2012), who considers meter as the cognitive organization of musical rhythms. I will return to this concept in subsequent chapters.
independently of us, and of a ‘time sense’ whereby a person perceives this
time. [. . .] It is useless to measure the sense of time against a clock which
is alleged to keep real time” (51). To soften the blow of Clifton’s critique, we
should note that there is nothing disagreeable about objective time in and of
itself, but only if it remains a tool for measuring the durations of our forever-
changing experience. Problems arise when one draws on the notion of ob-
jective time to conceptualize the temporal dimension of music, because this
forces musical processes to become reified as objects. They become associ-
ated with bounded, measurable “regions” of time, and placed in discrete tem-
poral “locations.” In turn, this gives form to the sound prior to the process of
that sound’s becoming, as willful a disregard of Hasty’s rebuke as there ever
was. Such reification is not necessarily questionable when it comes to trans-
mitting performance directions to musicians through visual representations
on the printed page—after all, even Norman’s and Frey’s pieces exist as phys-
ical scores—but it does close off the possibility of recognizing new and dif-
ferent kinds of meaningful experiences of this music. In particular, reification
conceals musical processes because objective time is only useful if there is
something “there” to be objectively measured. Yet, as coextensive with time,
musical processes could not possibly be “there” (or anywhere) prior to being
constituted by the listener who is in the midst of their becoming.

**Lived Time**

Instead of drawing on notions of objective time and the concomitant idea
that music unfolds within it, in light of the previous discussion, I propose
that musical processes are constituted by what the phenomenological tra-
dition has called “lived time” (Hoy 2009). Notwithstanding claims made by
models derived from physics, the sense of temporal passage is persistent,
even without any apparent change. We do not experience mere sequences of
events, series of isolated “nows.” We experience the processes of transi-
tion from one event to the next, the present, or what Maurice Merleau-Ponty
(2012) calls an “upsurge” contextualized by what came immediately before
and what is about to come after.26 This upsurge is not a durationless point
that separates the past from the future, but a kind of quality in our experience,

26 Komarine Romdenh-Romluc (2010) offers a wonderfully straightforward explication of
Merleau-Ponty’s (often opaque) ideas.
one that has both thickness and breadth. And, because of this contextual thickness and breadth, rather than basic relations of “before” and “after,” we actually have a sense of tense—of the past and the future. Events differ for us in terms of significance—in terms of how we position ourselves toward them physically and affectively—depending on whether they have already happened, are happening now, or will happen sometime in the future. Thus, the upsurge both frames and is framed by the specifics of the situation sensed from the perspective of an individual. There is something that pushes us from this present moment to the next one, and the one after that, and so on.

Lived time is time as it shows up in human lives. It does not so much stand in opposition to objective time, but has been argued to constitute its very source (Bergson 1911; Hoy 2009). It is easily conflated with subjective time—with which it shares a number of characteristics, including first-person experience—but the two are radically different. Subjective time refers to “the experience of the temporal properties of events and processes: their order, duration, time of occurrence, context among simultaneous events and events before and after, and more” (Arstila and Lloyd 2014, x). Since it is typically ascertained through judgments of its passage and estimates of duration, subjective time is basically construed as objective time from the perspective of the subject, hence internally represented and inaccessible to direct observation. By contrast, lived time is external to the subject but partly co-constituted by the subject’s interactions with the world. Subjective time paradoxically unfolds independently of the subject; lived time is the subject in unity with the world (Romdenh-Rayluc 2010). In fact, it is inaccurate to say that lived time is something that can be experienced as such, because experience requires an external frame of reference, which leads to models of temporality that are internally incoherent and confused (Merleau-Ponty 2012). Rather, lived time is enacted—it only exists as part of the unfolding dynamical system that emerges between an embodied consciousness and the world.

Lived time is a real aspect of the world in the sense of both describing and consisting of relationships between physical bodies and objects. In Chapter 4, I will take a detailed look at its structure, focusing on the role of human bodies in constituting it as part of what Merleau-Ponty (1968) calls

---

27 Merleau-Ponty, as well as Husserl, Heidegger, and other phenomenologists, were influenced by William James’s idea, discussed in his Principles of Psychology (1918), of the “specious present” as a temporally extended moment in which the subject has a sense of things happening “now.” For historical context see Andersen and Grush (2009).

28 Bergson (1911) coined the term durée to designate lived time.
“flesh.” There, I will show that time is “secreted” by a body that breaks away from its intertwining with its milieu. For now, let me briefly sketch some of its features.

Where subjective time is concerned with the order of our experience of time—and hence its quantitative aspects—lived time takes as its starting point the fact that the experience itself is temporal. It has a duration, but, as Edmund Husserl argued, this duration is not a measurement (Husserl 1964). Instead, it is a particular quality, a kind of intentional experience. By “intentionality” Husserl means our consciousness of something, or, rather, how our consciousness is always directed at something. Starting with a melody—Husserl's favorite example of a temporal object—he notes that its succession of tones is “united ‘forthwith’ in a common structure. [. . .] We do not have the sounds all at once, as it were, and we do not hear the melody by virtue of the circumstance that the earlier tones endure with the last. Rather, the tones build up a successive unity with a common effect” (Husserl 1964, 29).

The effect is puzzling if we maintain that we can only hear what is currently present, and the present can only include one note at a time, because under these conditions one would never be able to hear a melody as a melody but only as a succession of single tones. However, according to Husserl, perception itself has an indeterminate duration constituted by three phases. At the center of this structure, what is given to consciousness is the primal impression [Urimpression], which is our awareness of an unfolding “now-phase.”

Trailing it like a wake and grasping onto the immediate past is the retention, which is a quality of an event “sinking” into the past. It is a running off of a temporal phenomenon that allows us to perceive its continuity, and

---

29 An excellent introduction to the philosophical and historical developments regarding lived time can be found in Hoy (2009) and Andersen and Grush (2009), respectively. Lochhead (1982) provides an accessible précis of Husserl’s phenomenology of time consciousness, while Ihde (2007) offers a close phenomenological description of sonic experience in general. Godøy (2010b) explicitly links Husserl’s insights with bodily gestures.

30 The puzzle has its roots in Augustine’s well-known analysis of time, in which what he calls the distentio of the mind—a kind of stretching or pulling in different directions—contains within it images of the past and the future. As Gallagher (1998) notes, this creates a paradox, because events that ought to succeed each other in experience—such as the tones of a melody—are present to consciousness simultaneously. In Husserl’s solution, the extended consciousness both integrates and segregates the events that constitute the melody.

31 Husserl continued to develop these ideas over the course of his life, and for this reason the picture of his theory of time consciousness is dauntingly complex, particularly as concerns terminology. As explicated by Rodemeyer (2006), the phase of the primal impression was initially called a “now-point,” and this is the usage we find, for example, in Godøy (2010b).
a continuity between it and successive phenomena. Arranged symmetrically around the primal impression is the *protention*, a vague anticipation that reaches into the immediate future, an awareness that a phenomenon will endure or be replaced by a different phenomenon. Returning to the melody example, retentions and protentions “fuse with the apprehension of the tone that is now appearing and that, as it were, I am now hearing” (Husserl 1964).

Although at times Husserl referred to retention as “primary memory,” it should not be confused with memory proper, which he called “secondary memory” or “re-presentation.” An important distinction is that retentions are constitutive of the present, while secondary memories themselves have their own retentions. Similarly, protention is not a projection of one’s consciousness into the future; such a projection has its own tripartite structure. Rather than being of the past and the future, retentions and protentions make the experience of the past and the future possible. That is, we ourselves create the possibility of being temporal beings: beings with a history, with a destiny, with a temporal thickness.

Setting aside the problems with Husserl’s perception of a melody—including his treatment of it as a temporal “object,” which I will address in Chapter 4—his central claim is that our temporal experience does not unravel like a string of pearls, in which each pearl is discrete and self-contained, and yet exactly like every other pearl (Hoy 2009). Rather, each moment holds on to the one just passed, and anticipates the one about to come, as if each pearl seamlessly melded into the ones on either side of it while simultaneously being held as a discrete unit. The idea was treated explicitly by Merleau-Ponty (2012), whose reimagining of Husserl’s model of time

---

32 Gallagher (1998) points out that Husserl borrowed the idea of an extended present from William James (1918), who described the phenomenon as a “specious present.” Andersen and Grush (2009) have further shown that James himself appropriated the term from Robert Kelly, an American lawyer whom James knew through a friend.

33 For a cogent critique of Husserl’s experience of a melody, see Gallagher (1998).

34 Godøy (2010) draws parallels between Husserl’s consciousness of now-points and goal-points in sound-producing actions, arguing that they are analogous to keyframes in animation, while protentions and retentions are analogous to interframes. The problem with this approach is that it reinforces a string-of-pearls view of temporal consciousness, which is precisely what Husserl tried to refute. This may fall out of Husserl’s diagram, which shows primal impressions as points on a line, visually separated and strung up in a sequence, leading to the notion that consciousness moves from one now-point to the next through a fuzzy boundary of protentions and retentions. Instead, what Husserl argued (and Merleau-Ponty made even more explicit) is that consciousness moves in the primal impression, always flanked by the temporal fringes of the present. Godøy’s solution to this problem and a way of accounting for the continuous unfolding of behavior, is to introduce into actions the linguistic concept of co-articulation, in which movements intervening between goal-points are subsumed by impulse-driven holistic actions.
Meaning 37

Consciousness is shown here in Figure 1.2. Merleau-Ponty’s commentary clarifies some of the conceptual issues that have led to confusion regarding the nature of time. In particular, he argues that it is incorrect to suppose that time moves, or flows, in a succession of nows. “I do not pass,” he writes, “through a series of nows whose images I would preserve and that, placed end to end, would form a line” (439). Rather, time is “a network of intentionalities” (440). Instead of existing in succession, “instants” of time “differentiate themselves from each other” in a single, continuous phenomenon. “The springing forth of a new present does not provoke a piling up of the past and an upheaval of the future; rather, the new present is the passage from a future to the present

On the problem of Husserl’s static representation, his continued adjustments to the model in order to account for its dynamism, and suggestions of different types of diagrams that arguably do a better job of capturing it, see Larrabee (1994). On the relationship of Merleau-Ponty’s diagram to Husserl’s see Varela (1999), who notes in particular that there is no definitive diagram of Husserl’s model. I will have more to say about this figure in Chapter 4.
and of the previous present to the past—time sets itself in motion, from one end to the other, with a single movement” (442). Finally,

each present reaffirms the presence of the entire past that it drives away, and anticipates the presence of the entire future that is “to-come,” and that, by definition, the present is not locked within itself but transcends itself toward a future and toward a past. Thus, there is not one present and then another one that takes its place in being, nor is there even a present with some perspectives upon the past and upon the future followed by another present in which these perspectives would be overthrown, such that an identical spectator would be necessary to effect the synthesis of successive perspectives. Rather, there is a single time that confirms itself, that can bring nothing into existence without having already established it as present and as a past to come, and that establishes itself all at once. (444)

Lived time in this formulation is a single experience of a continually changing present, in which what was once implicit becomes explicit, while what was explicit becomes implicit. For Merleau-Ponty, driving this process is the subject’s “motor intentionality,” or a subjectivity directed toward the world through some kind of activity. For example, when I see my phone lying on the desk in front of me, I explicitly experience its top part, but I also implicitly experience its other surfaces that are currently hidden from me. As I turn the phone over, its implicit aspects become explicit (and vice versa), thereby linking the two experiences. Notably, my perceptual experience of the phone is not only spatial, but also temporal, because for Merleau-Ponty perception is guaranteed by the environment demanding some kind of action on my part. My motor intentionality in turn reveals a certain style, or a particular manner of performing actions, which gives coherence to the various episodes that constitute my life. In this model, there is no subject standing outside of time, gazing upon it as if it were a river, for, according to Merleau-Ponty, experience itself is a “plenum.” Since the plenum is all-encompassing, the past and the future are not separate existents, but instead constitute the present which “dehisces” toward the future. My actions convert the present simultaneously into the future and the past by responding to the possibilities solicited by the environment, by realizing their inherent potential. Hence, my subjectivity is “one single temporality which is engaged, from birth, in
making itself progressively explicit” (Merleau-Ponty 2012, 474). Lived time emerges from my union with the world.

***

The main concern of this book is precisely the lived time that I just described, and a discussion of how it is constituted will weave in and out of the main argument, particularly in Chapter 4. A central figure in my theoretical reflections is Merleau-Ponty, whose phenomenological approach to time explicitly involves the animate human body in its constitution. One of the key precursors to the field of embodied cognition, Merleau-Ponty argued that the most basic connection between ourselves and our world—a connection that serves as the ground of not just our experience but also our thoughts—is established through our bodies. In a commentary on Descartes’s notion of cogito, Merleau-Ponty (2012) claimed that “consciousness is originally not an ‘I think that,’ but rather an ‘I can’” (139). What is crucial to my argument is that bodies are actively engaged in “reckoning” with the environment, making sense of what they encounter by attempting to achieve an optimal perspective, or attitude (316). This achievement results in a particular kind of relationship that we have toward space and time, a first-person “I am a living being” perspective that leads Merleau-Ponty to posit that we do not so much exist in space and time, but rather that we “inhabit” them (an idea to which I will return in Chapter 6), meaning that our knowledge of space and time is fundamentally embodied and grounded in our abilities to skillfully orient ourselves toward the world.

Our skillful orientation—our motor intentionality—generates a field of possibilities within which time emerges as a lived aspect of our experience. The body here serves as the source of this emergence, but it is important to note that, for Merleau-Ponty, this body by itself does not constitute time. Indeed, time for him cannot be constituted, because for it to be constituted would mean that “the series of possible relations according to the before and the after” were fully known in advance of their happening. “Constituted time belongs to space,” he writes, adding that it is “immobile where nothing passes by and nothing happens” (437). Instead, time is always in the process of being constituted: “it must not merely be; it must come about” (437). This process sees the embodied consciousness as both active and passive. By passivity Merleau-Ponty means that time is a given fact of the world over which we
have no control. In his own words: “I am not the author of time, any more than am I the author of my own heartbeats, nor am I the one who takes the initiative of temporalization; I did not choose to be born, but no matter what I do, once I am born, time flows through me” (451). By contrast, the active dimension of the process signifies that we ourselves shape time by the decisions that we make: by making what was implicit, explicit. Accordingly, “this springing forth of time is not a mere fact that I undergo; I can find in time a recourse against time itself, as happens in a decision that I commit to, or in an act of conceptual focusing. Time tears me away from what I was about to be, but simultaneously gives me the means of grasping myself from a distance and of actualizing myself as myself” (451).

This book is an elaboration and extension of Merleau-Ponty’s phenomenology of time into the realm of musical experience. A key aspect of my argument is the role that the body plays in realizing the potential inherent in environmental solicitations, which allows for the possibility of experiencing the corporeal, evanescent musical objects and processes that emerge as the temporal meaning of music. It is a meaning formed by a fundamental relationship between music and time, one in which the two are inseparable, co-present, and co-constitutive. This inseparability finds its locus in the embodied listener who enacts musical time while simultaneously creating musical objects. Such an interweaving of time and musical meaning consists at the level that Elizabeth Grosz describes as “a pre-discursive experience, experience before the overlay of reflection, the imposition of meta-experiential organization, its codification by reason, language, or knowledge” (Grosz 1993, 43). This suggests a shift in how musical meaning emerges—from signification, as an expression that triangulates sounds with external signifiers, to a more basic, bodily engagement with the world. It is here that musical time manifests itself—not as a condition of music, nor a vessel in which it unfolds, but as a particular form of this experience: its enactment.

Of course, the performative aspect of signifying already has strong temporal and embodied dimensions.36 Time is enacted through acts of speaking and writing as much as through musical performance. In fact, I will show in

---

36 On the temporal aspect of linguistic communication, see in particular Auer (2015). Performativity is a term closely associated with the work of J. L. Austin (1962), who theorized the existence of so-called performative utterances, in which the state of the world undergoes a change as the utterance is expressed (e.g., “I now pronounce you husband and wife”), as well as John Searle’s closely related speech-act theory (1969).
Chapter 4 that time is enacted through all of our doings, even when we are not doing much. But the very units of expression—concepts, images, signs—are themselves non-temporal. They are like building blocks, necessarily maintaining their stability in order to facilitate symbolic communication. Certain bodily expressions, such as hand gestures that are used in communicative acts, function similarly, acquiring symbolic meaning when they gain ubiquity as repeated references to the same phenomena. This meaning can be conventional: “thumbs up” as a sign of affirmation,shrugging one’s shoulders to indicate nescience, or the cyclical hand gestures (kriya) that consistently mark the beats of the tala in South Indian classical music. But it can also be contextual, emerging ad hoc in the process of an unfolding discussion (D. McNeill 1992).

Other than these exceptions, the body does not disclose itself through a triangulation of meaning, but through action. Its mode of expression is analog, dynamically unfolding as its animate form encounters and makes sense of the world. Like the ephemeral musical objects discussed earlier, the body inhabits time, and in its unfolding it expresses an immanent meaning, which I call “significance.” This is a meaning that is pragmatic, non-referential, pre-discursive, and pre-reflective, consisting of the interactions between an embodied, situated listener and an acoustical signal. It can be modeled using J. J. Gibson’s notion of “affordances,” such that the very dynamics of this interaction are its meaning—its feel, its kinetic shape, its morphology (Gibson 1986, 1977). Significance does not point to an external interpretant, but rather is fully encapsulated by the totality of the situation and remains open to flux, continuity, as well as to temporal upsurge and disintegration. Time considered as significance becomes constituted as the form of the continually changing music-listener dyad.

37 Heidegger (1995), for example, talks at length about the link between temporality and profound boredom. For a discussion, see Hoy (2009).
38 Galen DeGraf’s (2018) dissertation includes a thorough discussion of the kriya as a dynamic temporal symbol that can be used to off-load temporal information, providing a time-dependent context, which helps listeners keep their “place” in the (sometimes very long) tala cycle.
39 Becvar, Hollan, and Hutchins (2005) present an ethnographic study of a fascinating example of a gesture that acquired contextual meaning and stabilized through ongoing discourse about molecular dynamics.
Significance

The idea that musical meaning lies in its significance is prefigured in the work of Susanne Langer. As she describes it, music is a “non-discursive logical form” whose meaning—or what she calls “vital import”—is constituted by feelings, emotions, motion, and life itself (Langer 1953). More specifically, the significance of music lies in its function as an “unconsummated symbol” for our emotive life, a symbol that lacks any fixed association but the form of which is nevertheless congruent with “the pattern, or logical form, of sentience” (27). Langer argues that “the basic concept is the articulate but non-discursive form having import without conventional reference, and therefore presenting itself not as a symbol in the ordinary sense, but as a ‘significant form,’ in which the factor of significance is not logically discriminated, but is felt as a quality rather than recognized as a function” (32).

Rejecting arguments that it is a stimulus meant to evoke or signal emotions, Langer contends that music is the logical expression of affects, by which she means that its structures “resemble certain dynamic patterns of human experience” (Langer 1957, 183).

The following quotation summarizes Langer’s perspective and foreshadows some of the themes that I will explore in later chapters:

The assignment of meanings is a shifting, kaleidoscopic play, probably below the threshold of consciousness, certainly outside the pale of discursive thinking. The imagination that responds to music is personal and associative and logical, tinged with affect, tinged with bodily rhythm, tinged with dream, but concerned with a wealth of formulations for its wealth of wordless knowledge, its whole knowledge of emotional and organic experience, of vital impulse, balance, conflict, the ways of living and dying and feeling. Because no assignment of meaning is conventional, none is permanent beyond the sound that passes; yet the brief association was a flash of understanding. The lasting effect is—like the first effect of speech on the development of the mind—to make things conceivable rather than to store up propositions. Not communication but insight is the gift of music; in very naïve phrase, a knowledge of “how feelings go.” (Langer 1957, 198; emphases in the original)

Two points are worth highlighting for their relevance to the discussion of enacted musical time. One is that musical meaning emerges below the level
of consciousness, in a pre-discursive domain of experience where understanding is shaped by affectivity and bodily engagement, rather than propositional logic. This meaning consists of patterns of movement, in the bodily orientation of listeners, and in their basic affective dispositions. It is a thoroughly temporal meaning insofar as it has no permanent features—it is instead an unfolding quality of the listening experience, an aspect of lived time. The listener’s affective posture—which includes certain corporeal dynamics, such as movement toward (philia) or away from (phobia) something—enfolds sonic forms within the temporalizing orbit of his or her body, essentially creating time in response to the acoustical signal. The second, related, point is that music presents an opportunity for our expression of a kind of knowledge of the dynamical patterns of emotions. In Chapter 3 we will hear echoes of this in the form of what I call “kinesthetic knowledge,” or a knowledge of what logically unfolding movements feel like, and how this feeling results from ongoing tension between habit and novelty. As I will show, the “insight” offered by music is one that arises from active responses to its affordances, more specifically, the temporal shapes enacted by an embodied, situated listener.

Despite these zones of convergence, Langer draws a sharp categorical distinction between art and artifact, where the latter—“like cake or cocktails”—appeals to the sensuous pleasure of “the untutored” (Langer 1957, 166). My position is therefore closer to John Dewey’s (1980), for whom art extends and augments everyday forms. In contrast to Langer, I see music as continuous with the ways in which our environment is significant for us: music emerges as one of the significant forms of our interactions with the world. To that end, in what follows I focus on significance as time, or, more specifically, as a temporalization of the world through our affective dispositions that guide our actions in response to its solicitations.

To illustrate the general principle of musical significance as an extension of everyday forms, let us first consider a somewhat prosaic example. Imagine that Mary is a high-school history teacher. She is sitting outside with a stack of final exams that need to be graded, when all of a sudden the wind picks up. At first just a breeze, Mary realizes that an unexpected gust would send her papers flying into the air. Visually scanning the space around her, turning her head this way and that, and shifting her body’s weight forward and back, she notices a rock, slightly larger than the size of her fist. She reaches out for it, and, feeling its heft, she deems it up to the task and places it atop the stack.
In this scenario, the rock attains a particular kind of significance for Mary. Although the signifier “rock” typically refers to a mass of hardened minerals, the concern here is not geological. Mary need not care about the exact chemical composition of this rock. It matters not how the mass Mary held was for her an example of rocks in general, or how it distinguished itself from other rocks, or how it fits into a larger narrative about the different ways in which rocks have been used by human societies. In fact, under the circumstances, it is entirely beside the point that what Mary was holding was something that, in English, is called “a rock.” This is not to say that all of these have not, at some point, shaped how rocks show up in Mary’s experience, including their capacity to be handled in a particular way. Rather, all that mattered at the moment when the wind threatened Mary’s grading session was that this specific object fulfilled certain criteria for doing the job at hand: it had the optimal size and weight to hold down the papers, it had the optimal size and weight for her to lift it, and it was optimally placed within her reach.

Whereas the linguistic meaning of this mass of minerals is “a rock,” its significance was precisely the action that unfolded in those circumstances. At that particular moment, and in that particular place, it was significant for Mary as something that she could lift and use to hold down her papers. The meaning of this mass of minerals was seemingly inherent in it; more precisely, it was inherent in the entirety of the situation, which included the outdoor setting, Mary’s own motivations, and her actions. In short, its meaning consisted of and persisted in the system that briefly consisted of Mary and her surrounding environment.

Meaning considered as significance is a practical meaning. It is the meaning that arises in the moment of one’s perception of—and action upon—one’s immediate environment. It refers to relevance, importance, and the capacity of objects and events to make a difference, as well as the organism’s ability to perceive these without having to “cognitively enrich” the world by projecting onto it structure and purpose (Wilson and Golonka 2013). Significance is expressed in bodily terms as an enactment of the potential inherent in the situation. It cannot be known in advance of the interaction that materializes it. To use the phrase made famous by Gregory Bateson, it is the “difference that makes a difference,” or a framework that allows us to pick out and use only that information which is relevant to the task at hand (Bateson 1985).40 It is precisely what enables someone like, say, Sherlock Holmes to take one glance

40 The phrase appears somewhat earlier in Johnson (1946).
at a room containing a dead body and almost immediately deduce the exact occupation of the deceased, where they lived, what they had for breakfast, and which train they took to get into the city. Yet the extraordinary capabilities of the fictional detective are a literary extension of our own remarkable powers of dealing with a complex world in ways that are fluid, automatic, and advantageous to our survival. We may not be able to glean, based solely on someone’s freshly shaved face, whether the light in their bathroom is on the right or on the left, but our world is nonetheless brimming with potentials for meaningful engagement. Significance emphasizes and draws attention to what is relevant and important in the environment. Instead of emerging from an interplay between an object and its external signifier, significance refers to our ability to use and respond to environmental solicitations in ways that can be called “relevant” or “appropriate.” Of course, whether and how our responses are relevant or appropriate varies with the situation and our objectives, but regardless of these variations, meaning is inherent in how we use and respond to the world’s solicitations.  

**Affordances**

Talk of objects being right for the job “at hand” might bring to mind Heidegger’s concept of *Zuhandenheit*; however, Gibson’s theory of ecological perception captures even more precisely the notion that meaning consists in a system that emerges when organisms interact with their environments (Gibson 1986). Having initially studied the landing techniques of bees and aircraft pilots, Gibson asserted that all animals access meaning in their environment through the direct perception of various opportunities for action, which he called “affordances.” According to him, meaning is inherent and available to the animal in the world that it inhabits: “the meaning or value of a thing consists of what it affords” (Gibson 1982, 407).

---

41 My “significance” is similar to Ian Cross’s “efficacy” in that both point to the individual and social function of music (Cross 2005).

42 The recent years have marked a surge in the term’s usage in music scholarship. These include Clarke (2005), Godøy (2010a), Krueger (2014), Menin and Schiavio (2012), and Windsor and de Bézenac (2012). Affordances also play a key role in the construction of musical meaning in Cook (2001) and London (2012). On the role of affordances in structuring music by constraining the relationship between musicians and their instruments, see De Souza (2017). Here, I limit myself to more general observations. (A detailed discussion of specific musical approaches will be taken up in Chapter 2.)
More precisely, the meaning and value of things is a factor of the relationship between variable and invariant features of the environment and the capacities of an organism to perceive and act on those features. Since the environment is already meaningful, Gibson’s radical position was that animals do not have to create internal maps, mental images, or schemata; calculate distances between objects; or process “data” apart from the information offered to them directly by the environment. In contrast to the theory of indirect perception—which holds that objects and events are not meaningful in themselves, but rather that meaning is created and stored internally in the form of mental representations—Gibson argued that animals detect “specific combination(s) of the properties of [an object’s or event’s] substance and its surfaces taken with reference to the animal” (Gibson 1977, 67). On the one hand, affordances consist of the elements already available in the environment, rather than existing as mental constructs that are projected onto the world. In this sense they are objective: as long as some organism with the right abilities exists, affordances will also exist, even when that organism is not physically present at that moment. To return to the rock example, the mass of the object and its size afford picking up and using as a paperweight even when there is no one around. On the other hand, affordances are not merely properties or features of the world taken by itself, and, indeed, not all objective properties can be construed as affordances. Instead, affordances are relations between the environment and a perceiver, and their purpose is to guide the actions of the latter. They are thus dependent on the perceiver’s capabilities, motivations, goals, and needs. That is to say that affordances do not inhere in either the environment or the perceiving organism considered by themselves, but rather in the system that emerges when the two interact. The rock described earlier does not afford picking up and using as a paper weight to an organism that cannot lift it, nor to one that has no need for holding things down to prevent them from flying off.

Although this aspect of Gibson’s theory will not figure much in this book, the idea that perception occurs without intervening mental representations initially had an important ideological motivation for Gibson. In particular, he was horrified by the scientific community’s reluctance to emphatically condemn Nazi atrocities, a reluctance that stemmed from an unflappable belief that all perception is culturally relative. Since moral arguments, on this view, are not fixed universals but rather products of social context, the utter barbarity of Hitler and his cohort could have been shrugged off as mere “inevitable irrationality.” For Gibson, this was reprehensible, and hence his campaign to ground perception in natural laws. See more in Costall and Still (1989).

Perhaps unsurprisingly, Gibson’s stipulation that perception is a process that takes place without mental representations has proven onerous to satisfy, because it requires a radical rethinking of the relationship between organisms and their milieu. Readers interested in the details of this rethinking might wish to turn to Golonka and Wilson (2016), Hutto (2013), and Withagen and van der Kamp (2010).
In Chapter 2 I will offer a more detailed glimpse of the theory of affordances in its application to music, but several key features bear explicating here in connection with significance. The first is that affordances are picked up by the entire perceptual system, which includes not just the collections of receptor cells attuned to specific kinds of energy and the brain areas to which they are connected, but also “parts of the organism that adjust, modify, or orient the receptors in active exploration” (Chemero 2009, 159). On this view, affordances emerge as combinations of changes and invariants in the perceptual array when the animal moves purposefully within its environment. There are two important points that have to be considered in this regard. One is that the body in its entirety is immediately implicated in the perceptual process, creating the very circumstances in which this process is able to take place. This means that the kind of information that is available to us is conditioned not only by the biomechanics of our bodies, but also by the social and cultural pressures that shape how these bodies are used in both day-to-day and aesthetic activities.

The second point is that musical significance—as much as it depends on our ability to engage levels of cognitive processing that include memory, abstract thought, and aesthetic judgement—is fundamentally grounded in our bodily skills, and how those skills are deployed depends on the specific context of the situation in which we find ourselves. In particular, we use our bodies in order to gather information about our environment. As Eric Clarke writes, “perception is essentially exploratory, seeking out sources of stimulation in order to discover more about the environment” (2005, 19). We move around in order to attain what Hubert and Stuart Dreyfus call a “maximum grip,” or a level of being embedded in the world optimized “in such a way as to bring the current situation closer to the agent’s sense of an optimal gestalt” (Dreyfus and Dreyfus 1999, 603). We move closer toward a painting in a gallery when we want to focus on the details, and away from it when taking in the whole image; we lean in when someone speaks softly, and step back from someone yelling; we make a fanlike motion with our hands to direct perfume droplets into our nose, and turn our heads when their concentration is overwhelming. Those movements create variance in the somatosensory array that constitutes our perception.44 They can be large—for example,

44 What I am calling significance is similar to Merleau-Ponty’s “primary meaning,” which refers to the way in which we get a grip on a situation by expressing structural similarities between our bodies and the intentional object (through rhythm, intensity, orientation, etc.). Primary meaning describes the determinate way in which we relate to something.
when you move your head and your whole body in search of something to use as a paperweight—or they can be minuscule. For example, Bompa and O’Regan (2006) have shown that even something like color perception, which we usually take to be completely automatic and immediate, is only possible because of the tiny movements made by our eyes. Without those movements we would not be able to see in color, this despite the anatomy of our optical system. Or, put differently, our anatomy evolved in tandem with our mobility, making them mutually co-dependent.

Affordances manifest in the actions that an organism performs in response to some aspect of its environment. Given this premise, significance is embodied, enacted, and situated. It starts with movement: with spatial and temporal displacement, with the change of posture and position, with the transfer of weight from one part of the body to another, with the shift in direction and force of muscular effort. It is inaugurated in the locking and unlocking of joints, and in the expansion and contraction of the body’s sense of presence within its “kinesphere.” When this movement connects us with the world, a kind of intimacy develops, such that there is no separation between ourselves and that world. For good or for ill, there is no disconnect that would sever our subjectivities from the objects of our perceptual acts. Through movement we inhabit the world. Significance embraces music’s proximity, a proximity that, while preventing it from pointing to anything concrete, allows it to depict what Merleau-Ponty describes as “certain schemata of Being, its ebb and flow, its growth, its upheavals, its turbulence” (1968, 123). The dynamic, mercurial, even precarious character of these schemata is what invites us to further investigate their nature and to examine living, fleshy bodies at their most vulnerable. As I will show in Chapter 2, musical meanings created in the process are grounded in the ways in which organized sounds temporalize the animate bodies of listeners.

Affordances are relations: musical significance emerges from the dyadic interaction between a listener and an acoustical signal that constitutes a very specific situation. An acoustical signal does not “mean” this or that.

45 The term “kinesphere” was developed by Rudolf Laban in Choreutics (1966) as a reference to “the sphere around the body whose periphery can be reached by easily extended limbs without stepping away from that place which is the point of support when standing on one foot” (10).

46 In characterizing affordances as relational properties of the environment-organism system I am following Anthony Chemero (2009). That said, there are dissenting claims that affordances are instead dispositional properties of the environment, and are complemented by the organism’s “effectivities” (Turvey 1992). More recent critiques of the relational view come from Sabrina Golonka and Andrew Wilson (2016). These critiques are substantiated by two important requirements that affordances must fulfill: they must persist in the absence of the organism, and the organism must be able to perceive how they engender behavior that is currently impossible (i.e., they must functionally support learning).
particular action, nor does an action “mean” this or that particular sound. There is no inevitability here, inasmuch as neither the signal nor the action represent, imply, presuppose, compel, or determine one another. Rather, actions and sounds unfold in tandem, each one responding to the other, together creating significance. Here I draw the reader’s attention to the distinction between response and reaction. While the latter term is often used to describe listeners’ actions stimulated by musical sounds, in reality it implies a causal link between those actions and a sonic stimulus. The stimulus always precedes the reaction as the logical antecedent for movement, even if in objective time the action anticipates the stimulus, as is the case with mean negative asynchrony (Repp and Su 2013). Response, by contrast, captures the idea that such actions are examples of cognitive behavior that has its own logical and rational organization, and which creatively complements the dynamical unfolding of the acoustical signal. As such, actions do not exist in a linear relationship with the external world, but rather form so-called dynamical systems with the components of the environment, in which temporal relations create a complex network that does not always causally correlate with objective time. As Chemero defines it in the context of cognition, “a dynamical system is a set of quantifiable variables changing continually, concurrently, and interdependently over time in accordance with dynamical laws that can, in principle, be described by some set of equations” (2009, 25). I will explore this idea at some length in Chapter 3, but for now it is sufficient to highlight the condition that dynamical systems are constituted by elements that are mutually dependent for their development, such that a constant of one element forms a variable in the second element.47

It makes sense to think that listeners respond to music with their bodies, because these responses are visible in such ubiquitous activities as dancing, foot-tapping, head-nodding, air-guitar-playing, and others. This joining of patterned sound and motor behavior exists in virtually every known culture around the world, and is responsible for what William McNeill (1997) calls “muscular bonding.” It forms an “emotional residue” that, through its evolutionary connection with the collective survival of our species, holds a powerful grip over human existence (W. McNeill 1997, 38). We see here a clear causal relationship, whereby the arrangement of sonic elements

---

47 An example that is often brought up is that of coupled oscillators, such as electrical systems or mass-spring systems. One of the most basic coupled oscillators is the Watt (or centrifugal) governor, which was used to control the speed of steam engines (Chemero 2009, 68ff).
Enacting Musical Time

exerts pressure on—and therefore imparts structure to—movements. But envisioning music as responding to listeners’ actions may, at first blush, seem counterintuitive: typically, music is thought to unfold independently of listeners. In contrast to a performer who creates musical structure, the listener is its passive receptor inasmuch as he or she cannot literally alter the acoustical signal produced by someone else. However, what I mean is that although an acoustical signal unfolds regardless of a listener’s actions—indeed, regardless of the listener’s presence—this presence is necessary if the signal is to be meaningful as music. To put it more sharply, only when there is a synthesis of an acoustical signal and a listener does music, as such, emerge. There must be a human being for whom this signal is “music” and not something else.

This view is an ecological fleshing out and reframing of Marion Guck’s premise that “music is created between some musical sounds and a person” (Guck 2006). She writes: “sounds do not become music until they have entered a person, until they have been heard or imagined and attended to. Music exists only in the interaction between sound and the body-and-mind of an individual” (Guck 1996). Both of these statements reflect a view that, far from simply absorbing sonic “information,” listeners essentially impart structure to the music in their perception, through their actions, thereby erasing the categorical distinctions between body, music, meaning, and even—as we will see later—affect. The Gibsonian perspective suggests that musical structure is the structure of the listener’s embodied engagement with the acoustical signal. Meaning is already evident at the level of bodily interactions. For example, tapping one’s finger to the beat—an action that is often relegated to basic, uncontrollable impulses of the undisciplined body—is already an act replete with meaning. Specifically, the meaning here is what, based on music-theoretical concepts, would be recognizable as “the beat,” but such a reference is unnecessary to seeing the tapping listener make manifest their understanding of the regularities that exist in the acoustical signal. In Chapter 3 I will show how this significance is thoroughly temporal, whereby the dynamics of unfolding sounds and the dynamics of a moving body coalesce and mutually co-constitute one another.

This characterization, of course, does not apply to someone who is listening to the stereo and turns the knob to change the volume, or someone who uses a computer program to alter the pitch of a recording. In these cases, the listener turns into the producer. Nor does it refer to the trivial (although by no means irrelevant—see, for example, works by Ryoji Ikeda) change caused by moving around the physical space within which the acoustical signal unfolds.
It is important to keep in mind that music in this case is not identical with, or equivalent to, the acoustical signal, and that musical properties may not necessarily be the same as acoustical ones. In particular, the acoustical signal is the pattern of air pressure waves that humans register as sound. In addition to the signal, “music” on this view also encompasses the listener, as well as the situation in which this signal unfolds, including all of its socio-cultural, political, and historical elements. Thus, whereas an acoustical signal has a quantifiable structure of its own, and this structure is furnished for the listener in the form of affordances, musical structure is a product of the system taken as a whole. This distinction is crucial, because it emphasizes the fact that music is as much a cultural product as it is an environmental phenomenon, and so its significance does not lie solely in its ecological import. As I will illustrate in Chapter 2, this artificial nature of music presents a challenge to current theories of musical affordances, urging us to reexamine with renewed precision what music affords for listeners, and how it does so. My position is that through its dynamical patterns it furnishes us with what I call “temporal affordances,” or cues for when to perform an action, which interact with very specific “social affordances,” or cues that govern engagement with other listeners. Later I will argue that these actions participate in the emergence of musical time.

Still, although not purely objective features of the world, affordances are determined by those features to some extent. Hence, even if it is not in itself a carrier of musical meaning, the acoustical signal—what can be thought of as the material presence of sound—has measurable aspects that provide part of the structure that subsequently has the potential to become meaningful (for example, energy in specific bandwidths producing a sensation of timbre). While significance may not be inherent in the changes of air pressure that create periodic waves, it is a function of certain features of these changes, which justifies carrying out close readings of the material object in order to suggest the different ways in which it may become actualized as meaningful in the context of listening. This point is worth emphasizing because, whenever concerns about listeners’ capabilities and cultural conditions are foregrounded, all too often detailed, systematic structural analyses of music are abandoned in favor of more general observations.

This is not to deny the fact that the systems of measurement, as well as properties of the acoustical signal to which they are applied, are themselves purely intersubjective. Thus, what our instruments identify as “bandwidths” are only so because we have identified them as such.
Scholars of ecological perception who expounded on Gibson’s ideas point to a relevant distinction between affordances and solicitations (Dreyfus 2014). On the one hand, I already noted that the rock-as-paperweight has many objectively quantifiable features, but few of them are relevant to the task of holding down a stack of exams. Thus, the rock’s size and weight are important in determining its suitability, but the details of its mineral composition—while certainly a factor in how weight is perceived—are of no immediate concern. On the other hand, affordances are defined as relations between features of the environment and the organism’s abilities. As such, they are not guaranteed to become manifest each and every time the animal comes into contact with a particular object. For a variety of reasons, abilities are subject to failure. For example, the rock-as-paperweight, despite its affordance for picking up, might slip out of Mary’s hand; or, trying to reach it, she might lean over too far, lose her balance, and tip over. The point is that Mary’s general ability to perform an action does not mean that she will always be able to perform it successfully. Furthermore, Mary may be variously unmotivated to perform the action, such as when she has no need for a paperweight—even if the rock, with its affordances, is present in her environment. The existence of an affordance alone does not mean that it will always show up in perception. Rather, some affordances turn into what Merleau-Ponty (2012) calls “solicitations.” With regard to musical significance, the fact that such a transformation takes place suggests that only a small part of all possible meanings ever becomes actualized, and it is precisely this activity of actualization that is fundamental to the enactment of musical time.

Affordances become solicitations through the body’s affective disposition toward its environment—something that I will examine in more detail in Chapter 2, and apply to music analysis in Chapter 5. Affectivity is what drives perception by motivating the animal to engage with the external world. Perception is affective insofar as it involves a sensitivity to the animal’s particular situation, which is always dynamic and which needs to be taken into account as a background against which experience unfolds. The stance which we have toward the significant meaning inherent in the affordances around us is rarely neutral, and instead falls somewhere on the spectrum between a positive disposition and a negative one. There is always an “urgency”—to use Silvan Tomkins’s term—in the way things show up in perception, and it is precisely this urgency that is expressed as affectivity.50 Affect, according to

50 Silvan Tomkins’s theory appears in a large number of his writings. Its most concise version appears in his essay “Affect Theory” (1984).
Tomkins, “amplifies” the world in a way that makes things matter to us, such that we are surrounded by objects and beings that hold for us some kind of value. Affects are “complex patterns of bodily reactions whose (biological) function is to respond to situations of urgent concern” (Hufendiek 2017). Given the choice of several affordances, affect is what enables the animal to pick out which one to engage with. Through this engagement, cognition and affect are intertwined with regard to the biological value of affordances. Cognition is thus not simply embodied—that is, situated in and constrained by the physical and physiological aspects of the body embedded in some environment—but it is also enactive, meaning that it is driven by motivations and concerns enabled by affectivity. Gallagher and Bower (2013) offer a lucid explanation:

Meaningful encounters with the world imply a perceiving agent with some basic motivation to perceptually engage her surroundings. Schemata of sensorimotor contingencies give an agent the how of perception, a tacit knowledge of potential sensorimotor engagements, without giving its why, which depends on latent valences that push or pull in one direction or another for attention and for potential sensorimotor engagement, reflecting, for example, a degree of desirability. (234)

Perhaps the most important point of drawing on the theory of affordances is that musical time is not a given. It is not an a priori form that exists independently of its contents or outside of the purview of an embodied, situated listener. Music is not meaningful in and through time, but in and through the dynamical system that forms when acoustical phenomena elicit responses from enculturated listeners that make these phenomena musical—that is to say, when the significance of sound is music. Music in this view is not an object that exists apart from the interaction between sounds and listeners. In fact, rather than an object, it might be better characterized as what Judy Lochhead calls a “network of sounding possibilities (tendencies)” (2016, 96). Time in this case is not a condition of music, but something that emerges from it, from within the network of listener-sound interactions. Rather than music depending on time for its meaning, music means time—a claim that I will elaborate in Chapter 2, where I posit that music furnishes listeners with specific temporal affordances set against a background of cultural expectations, and that it is at this level of the affordance system that musical meaning and musical time are inextricable, mutually co-constitutive, and dynamically implicated in one another.
Considering the issue of musical time in such a way presents us with an opportunity to situate it within embodied and enactive approaches to human cognition. This recontextualization not only enables us to use the general theoretical and methodological foundations that these approaches already supply, but also provides us with specific benefits for theories of musical meaning. First, having a robust epistemological structure in place gives us resources for talking about subjective experience while avoiding the pitfalls of solipsism. A combination of phenomenological and empirical approaches grounds subjectivity in the context of broader human cognitive capacities, while allowing for the nuances of the deeply personal, idiosyncratic bodily engagement with the musical world to emerge. Combining first-person reflection with quantitative analyses of observable, hearable, touchable—indeed, feelable—phenomena allows us to stem critique that we are dealing with “merely” subjective experiences, which cannot be generalized and are therefore largely irrelevant from the perspective of theory. Second, taking advantage of an embodied-cognitive frame of reference implicates musical meaning within the more comprehensive human bodily capacities for meaning formation. Such a frame reveals how music is both a reflection of—and how it is different from—everyday dynamical patterns (Zbikowski 2002). Third, drawing on enaction and embodiment gives us the means for considering actions taken in response to musical sounds as sufficient for cognition, and thus constitutive of listeners’ musical understanding. This point is crucial because in trying to move away from linguistic explication we need to find alternative ways of grounding our observations, ways that take into account unique aspects of bodily experience without devolving into unbridled subjectivism. As I will show in Chapter 3, the mobile body provides just such a ground under this conception, because its cognitive behavior, as something enacted, is visible and available for both quantitative and qualitative analysis. Finally, since time is a critical component of the dynamical systems constituted by affordances, considering musical behavior as part of embodied cognition presents us with an opportunity to develop new tools for dealing with issues of temporality. In turn, new temporal experiences can emerge in response to the myriad ways in which contemporary composers organize musical materials, a point that will concern us throughout the remainder of this book.