Musical Sound and Spatial Perception: How Music Structures Our Sense of Space

Mark S. Saccomano

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

Abstract

Musical Sound and Spatial Perception: How Music Structures Our Sense of Space

Mark S. Saccomano

It is not uncommon to read claims of music's ability to affect our sense of time and its rate of passage. Indeed, such effects are often considered among the most distinctive and prized aspects of musical aesthetics. Yet when it comes to the similarly abstract notion of space and its manipulation by musical structures, theorists are generally silent. My dissertation addresses this gap in the literature and shows how music's spatial effects arise through an affective engagement with musical works.

In this study, I examine an eclectic selection of compositions to determine how the spaces we inhabit are transformed by the music we hear within them. Drawing on Maurice Merleau-Ponty's theory of embodied perception, as well as research on acoustics, sound studies, and media theory, I deploy an affective model of spatial perception—a model that links the sense of space with the moment-to-moment needs and desires of the perceiver—to explain how these musical modulations of space occur. My claim is that the manner in which the music solicits our engagement affects how we respond, which in turn affects what we perceive.

I begin by discussing the development of recording technology and how fixed media works deemed "spatial music" reinforce a particular conception of space as an empty container in which sound sources are arrayed in specific locations relative to a fixed listening position. After showing how innovative studio techniques have been used to unsettle this conventional spatial configuration, I then discuss examples of Renaissance vocal music, instrumental chamber music, and 20th century electronic music in order to develop a richer understanding of the range of spatial interactions that musical textures and timbres can provide. In my final chapter, I draw upon these varieties of affective engagement to construct a hermeneutic analysis of the spatial experience afforded by Steve Reich's *Electric Counterpoint*, thereby modeling a phenomenological method for grounding interpretation in embodied, rather than strictly discursive, practices. By soliciting movement through the call for bodily action, music allows us an opportunity to fit together one world of possibilities with another, thereby providing an occasion for grasping new meanings presented through the work. The spatial aspect of music, therefore, does not consist in merely recognizing an environmental setting populated by individual sound sources. Through the embodied practices of music perception and the malleability of space they reveal, we are afforded an opportunity to reshape our understanding of the world around us.

Table of Contents

List of Figures

Acknowledgments

I would like to express my thanks to the faculty of the Columbia Department of Music for an extraordinary academic experience. I have received encouragement and inspiration from professors in music theory, historical musicology, and ethnomusicology, as well as from faculty in the Columbia Sound Art program. I would like to thank my advisor, Mariusz Kozak whose careful readings and clear advice were invaluable, and without which I would not have been able to navigate my way through such a challenging topic. I am also grateful to Joseph Dubiel for his interest in my work since I began my studies here, and the extra time he took to help me find my footing in a new discipline. I continue to be inspired by his attention to how we write about music and the importance he accords to careful written expression as a valuable tool in music theory. I would especially like to acknowledge Marion Guck, whose enthusiasm for this project from the beginning provided me with motivation to persevere on those occasions when I felt like I had hit a dead end and could not see the way forward.

Thanks also to Gabriela Kumar, who provided assistance on countless matters, both large and small, promptly and kindly. She has the remarkable ability to make students feel like VIPs.

To my fellow students in the department I owe a tremendous debt of gratitude. I never expected such warm support from so many of my younger colleagues. Returning to school after a long absence proved to be more challenging than I had anticipated, and I am afraid to think how I would have fared without them. I am fortunate to have had Eamonn Bell and Marc Hannaford in my cohort, who have always been ready to help and advise

whenever necessary, and always with openness and good humor. I also was lucky to have begun my studies in the same year as Russell O'Rourke and John Glasenapp. Their friendship helped make the intense early years of coursework and exams more manageable, and our latter successes more enjoyable. The music theory students who arrived in the doctoral program before me also provided copious support, advice, and camaraderie, especially Orit Hilewicz, Ben Hansberry, Maeve Sterbenz, Galen DeGraf, and Will Mason. Among the students in later cohorts, in both the music theory and the historical musicology programs, the friendships of Michael Weinstein-Reiman, Velia Ivanova, Julia Hamilton, and Mike Ford have been particularly sustaining. In truth, I have enjoyed working, and studying alongside all the department's graduate students; it seems unfair to only name a few here, as they all have contributed to my professional growth and personal well-being. My great thanks to each of them. Opportunities to be among so many bright, talented and generous people are very, very rare and I am both happy and humbled to have been a part of this select group of scholars for the better part of the last decade.

I also want to thank my entire family for their enthusiastic support during this long endeavor. Thanks to Dan Tirtawinata and Glenn Meyer for their encouragement, optimism, and reassurance over the course of decades. Tage Lilja and Kevin McAteer provided warmth, perspective, and, importantly, refuge, which helped refresh my reserves and keep up my momentum. And a very special thanks to Mikhail Angelovskiy, who provided encouragement on a daily basis.

PART I

Chapter 1: Music and the Contents of Space

1.1 Introduction: *Gesang der Jünglinge* as an Electroacoustic Monument of Spatial Music

From the first few seconds of Karlheinz Stockhausen's *Gesang der Jünglinge* (1955–1956) listeners are treated to a dizzying assortment of sounds coming at them from all sides.[1] The opening gesture of the piece consists of streams of gurgling electronic blips that rapidly sweep downwards in pitch before proceeding back up as their volume diminishes. An additional electronic sound, not unlike the flare of a suspended cymbal, accompanies the blips to herald the beginning of the piece. As the blips proceed to follow one another in rapid succession, like a stream of air bubbles released into the water, their location appears to shift: whereas at first they seemed to arise from positions to the left, right, and center, they soon move to one side, gathering together on the right, as new sounds enter on the left. Shortly after the sweeping down-and-up gesture, and the rapid swell and ebb of the electronic white-noise "cymbal," distinct pitches materialize—piano-like if not actual piano: E♭6 up to a very soft G6, then shifting down an octave for a melodic tritone, G5 to D♭6. Shortly thereafter an acoustic sound appears—a voice—at first mixed with another sound of similar pitch and timbre, then emerging clearly singing a simple two-note figure: a descending perfect fourth, E♭5 to B♭4.

Though the electronic, synthesized sounds of the piece are familiar to us now (and perhaps even quaint) they still harbor something alien and peculiar within them. The electronic sounds are, to be sure, delivered though accessible gestures— a fall and rise in

1 Available on Spotify: https://open.spotify.com/track/0Y9DYv05QW8Cs8UTrXnRnA?si=9SOI61j-Say6BOE_ivPweA.

pitch, identifiable intervals, and, without much delay, the accompaniment of a human, singing voice—but the fact that these sounds cannot easily be imitated by familiar acoustic instruments provides this introductory passage with a remoteness that is often characteristic of electroacoustic music. This opening also contains a display of what was (for Stockhausen, at least) a defining element of this composition: the panning of audio signals across multiple channels, played back through a corresponding set of loudspeakers dispersed through the performance venue.[2] In this way, the sounds recorded onto the tape were heard to be coming at the listener from more than one direction, and could even be heard to change position from one side of the listener to another, leaving them unmoored, floating, as opposed to what would typically be heard in a stereo recording of a small ensemble, for example, where identifiable instrument sources stay put. Sounds with sources that could not be seen and with sources that were difficult to imagine could nonetheless be perceived as being in motion, shifting location from one place to another within the listener's environment.

The opening series of pulses that make up the swooping down-up gesture trace a movement from the center of the sound stage to the right. ("Sound stage" is a term coined by sound engineer and electronic music scholar William Moylan. It is used in connection with recorded music to refer to "the area within which all sound sources are perceived as being located.")[3] As the sound progresses rightward, its volume decreases and artificial

2 Patrick Valiquet discusses how Stockhausen repeatedly claimed that *Gesang* was the first musical work to make use of the property of tones "being generated in a definite space and in definite locations in this space...the tone-location remained fixed for all music written until now." Stockhausen, "*Musik im Raum*" (1959, 64), cited in Patrick Valiquet, "The Spatialisation of Stereophony: Taking Positions in Post-War Electroacoustic Music," *International Review of the Aesthetics and Sociology of Music* 43, no. 2 (2012): 414–15.

3 William Moylan, *Understanding and Crafting the Mix: The Art of Recording* (Boston: Elsevier/Focal Press, 2007), 50. Moylan notes elsewhere that the sound stage is "a two-dimensional area (horizontal plane and distance) where the [recorded] performance is occurring." Moylan, 178.

reverb is mixed in, causing the individual sounds' onsets to blur into one another. The result is a stream of sound that not only moves to the right, but seems to continue in the same direction, reaching a point that seems far-off, its source out of reach. As the stream fades away, the piano-like tones and the entrance of the singer's voice are audible in a different location, on the left-hand side of this projected sound stage, distinct from the fading, gurgling electronic pulses.

Gesang der Jünglinge is a central work in the history of what is generally termed "spatial music." Its success and notoriety had the effect of locking in a particular conception of space for electroacoustic music, and moreover, of claiming that type of space as a distinguishing feature exclusive to its domain. As Patrick Valiquet explains, this stance was a strategic one:

> Spatialisation helped to carve out a new position for electroacoustic music within a broader field of musical production. An ideology of aesthetic isolation supported the otherwise contradictory work of appropriating the tools of commercial recording and broadcasting.[4]

Commercial recordings, driven by a concern with verisimilitude, were routinely fashioned to resemble stage settings by recreating room acoustics and instrument placement.[5] Electronic art music sought to repurpose and subvert those effects. Valiquet also describes how a newly institutionalized avant-garde was able to marshal the resources of new and complex electronic technology to serve their aesthetic ends. He notes the founding of NWDR (Nordwest Deutscher Radio Rundfunk) in 1945, the Darmstadt Summer Courses in 1946, and the RTF (Radiodiffusion/Télévision Française) in 1949 as the

4 Valiquet, "The Spatialisation of Stereophony," 415; Valiquet, 415.

5 The convention of mixing tracks to resemble the placement of instruments on a stage is described by Simon Frith, *Performing Rites: On the Value of Popular Music* (New York: Oxford University Press, 1996).

beginning of a collaborative network among artists and technicians working in the post-war Western art music tradition. Shortly thereafter, in 1952, the *Studio für elektronische Musik* was established at NWDR Cologne—renamed WDR in 1956. As Jennifer Iverson has recounted, WDR became the "central node" in the "coherence of the mid-century avant-garde."[6] Her research uncovers an intense collaboration between artists and technicians, as well as the "composers' interdependence upon each other" as they struggled to master the complex technology at their disposal. The existence of this network meant that Stockhausen's *Gesang der Jünglinge* had the potential to influence composers of electronic music for years to come. Stockhausen's innovation of combining the acoustic, "found" sounds characteristic of *musique concrète* with the electronically generated sounds of *elektronische Musik*, along with his composerly virtuosity in segmenting and recombining these elements, meant that *Gesang* became a genre-defining work for electroacoustic music.[7]

Stockhausen himself claimed *Gesang* to be revolutionary in its use of space for musical purposes.[8] It is clear from the piece itself, as well as from his published remarks, what space meant to him: the direction from which sounds originated, and the trajectory they appeared to take. In the four-track version of *Gesang*, sounds are heard as not only having a location, but often exhibiting movement as well, coming toward or receding from

6 Jennifer Iverson, "Invisible Collaboration: The Dawn and Evolution of Elektronische Musik," *Music Theory Spectrum* 39, no. 2 (2017): 202.

7 Moylan was able to state confidently in 1982 that "In E.M., [electronic music] pitch, duration, loudness, timbre, and the spatial properties of sound have an equal potential to be the primary generator of the musical material;" and additionally, "sound localization and sonic movement have become vital components of certain electronic works." William Moylan, "An Analytical System for Electronic Music" (Ball State University, 1983), ProQuest (303122930), https://search-proquest-com.ezproxy.cul.columbia.edu/pqdtglobal/docview/303122930/ AF6A649A6D2F4856PQ/1.

8 Karlheinz Stockhausen, "Electroacoustic Performance Practice," translated by Jerome Kohl, *Perspectives of New Music* 34, no. 1 (1996): 74–105, and others.

the listener, or moving about in a circular fashion along a horizontal plane.[9] Trajectories are also heard in the two-channel mix Stockhausen prepared for radio and commercial distribution: sounds appear to emanate from the right, center, and left of the listening field; often heard moving from one of these locations to another. In addition to lateral placement, Stockhausen was able to create an impression of a sound receding, by adding artificial reverb along with a reduction of amplitude.

Focused as this work was on the possibilities unleashed by audio technology, *Gesang* still retained a strong visual element. A photograph from the premiere performance at the WDR concert hall in Cologne shows a pair of enormous loudspeakers placed on stage with a sea of heads expectantly turned toward them (Figure 1). The careful stage design was no doubt intentional: it provided a conventional framework within which to present a highly non-conventional, unfamiliar work. Yet this staging was also necessitated by Stockhausen's specifications for *Gesang*'s performance. For sounds to actually appear to arrive from the intended locations inscribed onto the magnetic tape, an audience was required to sit equidistant from four loudspeakers placed in the corners of the hall.[10] Though there would be variations from a presumed "sweet spot" in the center, general directions like left and right, front and back, could still be clearly perceived. (Such directional information would be radically misinterpreted were a listener seated facing the rear of the hall, or ninety degrees to either side, not to mention if audience members were allowed to wander about

9 Stockhausen's original conception of the piece included a fifth channel that was to be played from a loudspeaker suspended above the audience, a feature that was abandoned because of difficulties in installing the speaker in the performance space. Instead, the fifth speaker was placed on the center of the stage facing the audience. Later, the piece was mixed down to four channels, the standard version that is most commonly performed.

10 The loudspeakers were placed in the corners of the room as opposed to the cross formation of front-back and right-left speaker pairs that Stockhausen used for early performances of *Kontakte*. Simon Emmerson, *Living Electronic Music* (Burlington, VT: Ashgate, 2007), 156. Photo:

on their own.) Interestingly, while Stockhausen experimented with placing loudspeakers around the listeners, and engineering sounds to travel from one speaker to another, the position of the audience remained the same as it had been over the preceding 150 years: sounds were aimed at an immobilized audience that was subjected to their impact.

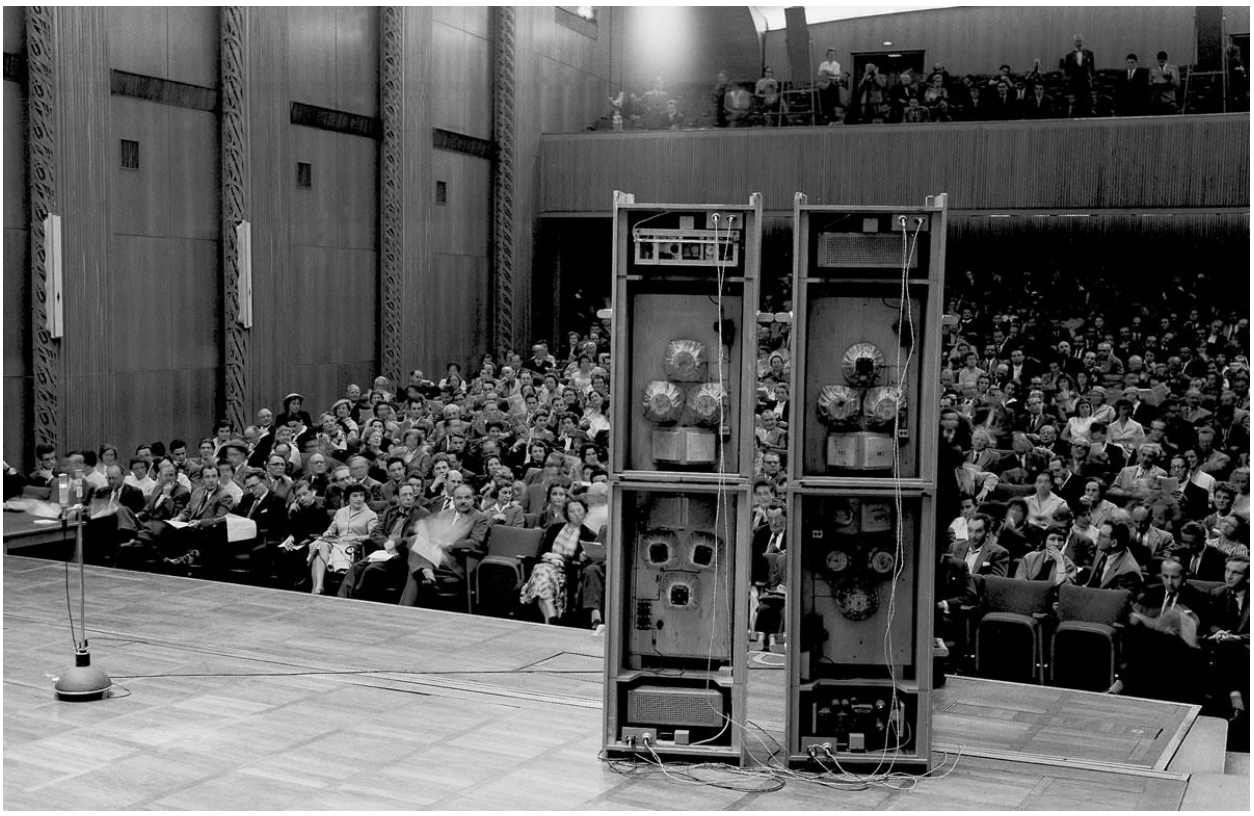

Figure 1: Premiere, *Gesang der Jünglinge*, May 30, 1956, Große Sendesaal, WDR Studios, Cologne.[11]

This performance setup also reflects a specific idea of what the term "spatial" refers to when applied to music. Though *Gesang* certainly represents a technical expansion of

11 Source: http://sites.music.columbia.edu/masterpieces/notes/stockhausen/GesangHistoryandAnalysis.pdf.

sonic possibilities, the space in the work is configured as an empty container, a stage populated with sounds as if they were actors in a play (though with the WDR's new playback system, the stage could encircle the listener). The use and reinforcement here of a long-observed norm for representing space recalls a remark by Simon Frith: "What is most startling about the history of twentieth-century sounds is not how much recording technology has changed music, but how little it has."[12]

In this theatrical scene of distinct sonic components, the directional attributes left, right, front, and back comprise an additional compositional element akin to articulation or dynamics. A new parameter is introduced allowing what was once a para-musical performance constraint to become part of the intra-musical compositional structure. Though not without its precursors, Stockhausen's rigorous and systematic deployment of this parameter, along with the newness of the technology, allow it to be rightly considered an innovation. But the intriguing element here, and one that rarely gets noticed, is that space is defined solely as location. That is, in *Gesang*, the spatial component of music is defined by the apparent location of the composition's constituent sound sources: space and location are treated as mutually implicative and mutually constitutive. If a sound object moves from one location to another, then there must be a space in which the sound is permitted to do so; if sounds are placed relative to one another, then there must be an absolute space that accommodates the represented locations. Simply put, if locations are indicated, then space is indicated as well. This is certainly a reasonable conclusion to make, but considered in this manner, space is solely a byproduct of the physical location of external objects, dependent upon the coordinates of the objects that call it into existence.

12 Frith, *Performing Rites*, 245.

Of course, there is nothing untoward about this conception of space in music, but it is a conception born out of a network of historical events, social structures, and individual activities resulting in a perceptual framework that supports this everyday idea of space. It is a model used not only in spatial music, but in other styles of Western music. In this chapter, I will interrogate the naturalness of this notion of space, and show that, even within such a framework, music provides access to other varieties of space as well.

First, I will discuss the idea of space as it is understood by composers and performers of spatial music. It is a conception of space rooted in both historical performance practices and technological developments in sound recording. I offer as an alternative a phenomenological understanding of space, one that is shaped by music. I then talk about framing as a property and a practice that provide cues for how we approach musical works. It is a term borrowed from theater studies, via sociology, into media theory. The idea of frame not only provides an understanding of the listener's role in a work's realization, but also how the recording and listening practices that have come to dominate our understanding of musical space evolved out of social practices. I discuss Mark B.N. Hansen's use of framing and his theory of virtual reality as "mixed reality," a concept well suited to describe the activity of listening to recorded music.

To illustrate this traditional notion of space reflected in recordings, as well as how the manipulation of conventions can open up different ways to understand space, I examine a pair of works by Miles Davis from the early 1970s. The tracks solicit a reframing of the music that incorporate the listening situation as well. These analyses lay the groundwork for my discussion in Part II of how music affects our spatial perception. There, I will discuss Maurice Merleau-Ponty's theory of embodied perception and show how the

texture and timbre of music can affect our sense of the space around us. Finally, I will show how these embodied engagements with music can serve as a ground for a hermeneutics of music that does not rely on a theory of musical representation.

1.2 Where is Space? Preliminary Remarks

The foregoing discussion of *Gesang der Jünglinge* exemplifies how this study will proceed. Rather than choosing in advance a definition of space by which to gauge individual musical examples, or attempting an amalgamation from various sources, I will draw out an understanding of space that the music presents to the listener. In *Gesang*, the marked use of directional cues in the audio channels, along with a performance configuration in part necessitated by the fixed-medium format, provides a topographical depiction of space as a set of identifiable locations marked out in sound and placed about a stationary subject. The same approach will be used for the other works in this dissertation: I will describe the manner in which the music solicits and draws upon a particular perception of space, and extrapolate the musical structures that provide for such an experience.

Space is an entity we reckon with in music all the time, whether we realize it or not. Though often passed over in music-theoretical research in favor of its kindred concept time, space is fundamental to how we approach music in any listening situation. The reason for this is that space is an integral component of perception: it functions as a realm that envelops both the listener and the music in a shared field that makes meaningful expression possible. Without an account of the role space plays in music, and how music affects the space we are in, our theories of music perception are apt to be limited to those which conceive of music listening as a transfer of information (with music relegated to the status of data). This is insufficient to explain one of music's most valued attributes: its effect on our moods and feelings.[13]

In Part I of this dissertation, I will examine the phenomenology of the listening event—in particular, the ubiquitous yet peculiar and still under-theorized musicking practice of home listening, an activity that exists ontologically somewhere between command performance, archival visit, and virtual reality amusement park ride. I will examine the behavior of sounds and the experience of listeners in the incongruous melding of the spatial cues inscribed in a stereo recording with the acoustic characteristics of the playback environment. Studies of the spatial qualities of recorded music typically focus on its engineered spatiality, that is, the echo, reverb, and stereo imaging present in the recording; few discuss the presence of these sonic features within the acoustics of the ultimate listening situation and how their commingling affects the spatial perceptions of the listener.[14] Instead, such studies generally consider the recorded spatial cues to supplant those of the listening space, a case of one acoustic environment being swapped out for another.

1.2.1 Replacing Space

An analysis by Eric Clarke provides an example of this tacit assumption that the acoustics of a recording supplant those of the listening space. In a spatial analysis of a pair

13 Tia DeNora describes how people use music as a tool to regulate their mood and provide both solace and motivation on a regular basis. Tia DeNora, *Music in Everyday Life* (New York: Cambridge University Press, 2000).

14 Simon Zagorski-Thomas (along with Peter Doyle, discussed below) is an exception here. He explains how dance records, which would likely be played in a discotheque—a large, cavernous space that would naturally add extensive reverb and muddy the pivotal rhythmic patterns in the bass register (bass guitar and bass drum)—were recorded and engineered to compensate for the loss of clarity. Simon Zagorski-Thomas, "The Stadium in Your Bedroom: Functional Staging, Authenticity and the Audience-Led Aesthetic in Record Production," *Popular Music* 29, no. 2 (2010): 251–66.

of popular music tracks, Clarke claims that musical sounds, "like other sounds, specify the various spaces from which they emanate."[15] He provides a guided tour through the landscape that he as a listener perceives: "an environment whose size and 'density and style of habitation' I experience through the spaces, textures and events of the music." Thus, he describes a certain mix as having the "slightly claustrophobic quality of a confined space" and the introduction of a new instrumental sound as providing "the spaciousness that comes from a moderate level of reverb."[16] He does not give an account of his actual listening situation and whether it has any impact on his environmental experience.

Peter Doyle is one of the few researchers who acknowledges the co-presence of competing spaces that arises with recordings. As will be discussed later in this chapter, he mentions the production of "descriptive records" in the 1890s, recordings intended to represent the sound world of a particular setting or environment, as with one such record titled *Morning on the Farm*: "This 'sound picture' overlays the fictive spatiality of the 'picturesque farm' onto the actual space of the 'parlor,' without challenging the integrity of either space."[17] And when considering spatial features of music, the parlor needs to be taken into consideration: it is important to remember that any relationship to the music is the result of the listener's mode of engagement rather than a quality that inheres solely in the recording or the performance or the work.

Movement and space are linked, and our possibilities for movement are determined by the spatial qualities of our environment. A piece invites us into the world it projects by

15 Eric F. Clarke, *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning* (New York: Oxford University Press, 2005), 109.
16 Clarke, 104–5.
17 Peter Doyle, *Echo and Reverb: Fabricating Space in Popular Music, 1900-1960* (Middletown, Conn: Wesleyan University Press, 2005), 50.

soliciting its own unique mode of listening. When music can affect how we perceive space, it can simultaneously affect what it is we are able to do. Maurice Merleau-Ponty, discussing visual art, writes how we are invited to take up the gesture offered by a work. This suggests that one place to look for the significance of expression is in the field of action and the spatial arena that accommodates the individual listener

As it happens, two types of sounds—the very simple and the very complex—prove to be excellent examples of how an awareness of space emerges in the activity of listening. Sounds without overtones do not reveal reliable information about the location of their sources. Sounds with multiple, chaotic, inharmonic overtones also confound our ability to determine their location. And if one sound masks another's reverberant characteristics, any discernable sound source locations in three dimensions become unstable. Using the acoustic characteristics of such sounds and taking embodied perception as a model for music listening, I will chart out a channel for meaning that opens up between the music and the listener that will help explain how music impacts the lives of its listeners and, moreover, how it can give us resources for reimagining the spaces we inhabit and the manner in which we inhabit them.

1.2.2 Methodology

The analytical approach used in this dissertation is primarily a listener-based phenomenology. Because I am interested in felt dimensions of space, I will attend to affective qualities I experience while listening to specific moments of a musical piece. Examining a score (when one is available) can only point us to a subset of possible affective listening experiences. We can speculate with some degree of confidence that a listener might feel surprised if we see a soft passage followed by one marked *fortissimo*, or make an educated guess regarding the degree of tension a listener might experience by noting the dissonance of a certain harmony in a given context. But scores do not carry information about individual performances and the varying conditions of different listening events. As such, the reliability, or repeatability, of the conclusions of phenomenological analyses should not be the sole criterion of their validity. Instead, I offer my analytic descriptions as plausible hearings, meant to serve as an illustration of possible effects of a composition's musical materials.

The fact that a sense of depth emerges for me through listening in at least one listening event may open the door to a way of describing effects not otherwise accessible to conventional music-theoretical approaches. As Eugenie Brinkema (2014) suggests, we can and ought to give a deep reading of these affective facets of aesthetic experience.[18] Otherwise, by avoiding them as too difficult and unruly, we end up reinforcing a positivistic ideology that the mainstream of humanities scholarship often tries to define itself against.

18 Eugenie Brinkema, *The Forms of the Affects* (Durham: Duke University Press, 2014), xvi.

To link these experiences with the sonic materials, I turn to philosophical and psychological works on embodied perception and scientific research on acoustics to show that such affective experiences are properly musical ones. That is, these experiences are linked to the music within particular listening events. Such links ground phenomenological analysis in a broader context that helps explain why certain musical structures affect us the way they do. Specifically, I will be investigating how sound waves generated through harmonic and inharmonic resonances interact within a listener's environment to create variations in pitch, timbre and texture that draw listeners' attention to their surroundings and their own place within them. These interactions implicate a moving body and I consider them to be inherently spatial. It is the purpose of this dissertation to draw out music's role in affecting their shape and quality.

My analyses investigate a specific variety of musical analysandum: listening to recorded music played through loudspeakers. And while their findings are not limited to a home listening experience, a domestic setting is in fact where these analyses were conducted. Headphone listening is alternately considered the ideal representation of a recording as sonic work, or the opposite, a delivery device that undoes the care put into a recording that has been engineered to be played through loudspeakers. Simon Emmerson and William Moylan rightly note that in either case, there is no control over the end listening situation, or the quality of speakers or headphones used.[19] More importantly, headphone listening is a different object of study—one in which the sound entering the ears is highly controlled. While headphone listening also features a layering of spatial

19 William Moylan, "The Elements of Sound and Audio Recording," in *Understanding and Crafting the Mix: The Art of Recording* (Boston: Elsevier/Focal Press, 2007), 147.

frames upon one another, the audio portion of the scene is stubbornly resistant to the moving body—a powerful effect and one and worthy of careful study. However, because I want to move away from studies that consider the acoustics of the recording to supplant those of the actual listening environment, I will be concerned exclusively with loudspeaker performances (and with any live performance situations that can be extrapolated from them). As each of the following analyses will demonstrate, an investigation into the infiltration of one space with another can provide important clues to music's affective force, not just in this particular musicking activity, but in a wide variety of performance and listening contexts where the spatial aspects of texture and timbre play an important role.

1.2.3 Space and Affect

I started this dissertation with an example of a fixed media work because our ideas about space in music have been thoroughly influenced by the birth and development of sound recording technology. Early uses of wax cylinders and later marketing of phonographs for the home defined listening practices for years to come, paving the way for the musical innovations that will be discussed in this chapter. But these ideas about musical space are thoroughly rooted in a long tradition of Western philosophical speculation on the nature of space in general. Philosopher Edward Casey describes the common concept of space as "impassive absoluteness," and attributes it to Plato.[20] He

20 Edward S. Casey, "How to Get from Space to Place in a Fairly Short Stretch of Time: Phenomenological Prolegomena," in *Senses of Place* (Santa Fe, N.M.: School of American Research Press, 1996), 13–52.

details how Plato conceived of space as a receptacle, a preexisting necessity in which the creation of the world took place.

> *Space:* a matrix for particular places that is ingredient in and coextensive with the Receptacle as a whole; to be placed herein is to be placed in Space *(chora),* that is, to be placed *somewhere*… in the Receptacle regarded as a massive spatial sphere, beyond which there is Nothing, not even the Void.[21]

This concept of space as a container for objects located within it—though eminently familiar to us—is not the only way to understand space. Music can provide a means of access to different ways of understanding our spatial relationship to our surroundings. (I will consider some of these alternatives in Part II.) We also make assessments of physical space in affective terms. For example, whether a particular locale seems open and expansive or closed and confining, is the result of a felt quality of the space, not wholly dependent upon measurements of absolute size. These felt qualities take our discussion of space beyond the "empty receptacle" model. The descriptions themselves demonstrate how our understanding of space is thoroughly imbued with affect and quickly moves beyond specifications in terms of metric quantities of breadth and width and height. A very simple thought experiment can demonstrate this fact: imagine walking into a room painted white, and compare that experience to walking into a room painted black. The latter would doubtless feel smaller. Similarly, a space can feel different when music is playing, a difference we might also describe, and in fact often do, in spatial terms.

This affective component is a necessary part of our perceptive apparatus. According to Maurice Merleau-Ponty in *Phenomenology of Perception*, perception necessarily involves

21 Edward S. Casey, *The Fate of Place: A Philosophical History* (Berkeley: University of California Press, 1997), 41. Italics in original.

the body and hence, an organism that reaches out to take in the world around it.[22] Grounded in the body, Merleau-Ponty's theory is thereby grounded in affect, as a perceiver's perspective is negotiated according to the concerns of the perceiver's body. The theory is thus one of *embodied* perception. Even for researchers who are not explicitly phenomenological, such as the psychologist and cognitive scientist James J. Gibson, perception is not a reflex that occurs in response to external sensations, but is instead an action, or achievement.[23] And it is an action undertaken according to the moment-to-moment wants and needs of the perceiving organism. Affect is not secondary to our perception, but bound up with our ability to perceive anything at all.

Much of the speculative and experimental research that surrounds affect considers it to be a pre-verbal feeling. This characteristic distinguishes affect from emotion, which is often considered to be a cognitive reflection on affective experience. Furthermore, what we feel is tied up with what we sense we are able to do in a given situation.[24] Affect therefore can be considered a way of being attentive: a readiness to respond that incorporates our bodily capabilities to act and to react. In addition, for Brian Massumi, it is an attentiveness that acknowledges ourselves and others as potential actors (or "actants") in a shared space—a space for *inter*action.

In *Parables for the Virtual* Massumi explains that affect is "intensity," and intensity, in turn, is the possibility of action in a given situation. He is careful to qualify the sort of

22 Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Donald A. Landes (New York: Routledge, 2012).
23 James J. Gibson, *The Ecological Approach to Visual Perception* (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1986), 148.
24 Brian Massumi, *Parables for the Virtual : Movement, Affect, Sensation* (Durham [N.C.]: Duke University Press, 2002); Scott L. Marratto, *The Intercorporeal Self: Merleau-Ponty on Subjectivity* (Albany: State University of New York Press, 2012).

action he is considering: an act that is not exactly premeditated, nor an unconscious reflex. To explain, he gives an example of a soccer match and describes how the boundary between subjectivity and objectivity is not clear in this context. Recalling the work of Bruno Latour and actor-network theory, Massumi notes that although a player may kick a ball, it is a mistake to consider the ball the object and the player the subject. The ball, he says, is a "part-subject," the organizing principle: "The ball arrays the teams around itself" and is the locus of change in both teams' scoring potential as its position moves about in relation to the goals at either end of the field. In this manner, "the ball moves the players." The player, then, becomes a "part-object" that is put into "a state of intense readiness for reflex response." What is affected by the player's kick, then, is not so much the ball as the state of the game.[25] Here, the conventionally understood categories of subject and object are replaced by varying propensities of affecting and being affected. A quality does not properly reside in any single experiencer. What players are attuned to moment to moment—the position of their bodies, the location of the ball, the state of the game—determines to a large extent what they are affected by. In this model, the phenomenon of being affected is not entirely a passive one. And even an inanimate object can become an affecting agent. Music, like the soccer ball, manages to be affecting in a similar manner, with a listener taking the role of the player, and experiencing the music as the "locus" of their own affects.

Affect emerges from this model as a relation. Our own selfhood is bound up with these engagements with the environment. An object of interest offers new opportunities for engagement—and our body senses that movement and change in position will provide

[25] Massumi, *Parables for the Virtual*, 73.

new perspectives: these are the possibilities for action that music affords. Through guiding direction of those movements and offering new perspectives, music alerts us to possibilities for ways of being, modes of behavior, and senses of self. This capacity of music to effect change is most clearly evident in the manner in which it solicits physical movements in space.

1.2.4 Space and Embodied Perception

Philosopher David Morris provides a clear and vivid example of how our perception of space is dependent upon the body and its concerns. "How far away are my glasses in the night?" he asks. "The distance is expressed for me in the movement of trying to envelop first my nightstand and then my glasses."[26] Here is a case where our understanding of space is not rooted in vision, of a depth that is defined by how objects occlude one another from different perspectives (this visual framework grounds much of Merleau-Ponty's discussion of space). As such, Morris's example clearly demonstrates how the sense of space is wholly dependent upon a motivated body with the ability to reach and grasp the things around it. Without these aspects of perception—intention and possibility—we are left only with an abstract concept of space as an emptiness that is measured by a three-dimensional grid. Uninhabitable, such an empty space is unimaginable. Though we may consider such emptiness measurable along perpendicular axes, a coordinate grid only provides a system to note position. It says nothing about what space itself *is*. It simply records where things are relative to an arbitrary set of boundaries. This is what leads Casey to the conclusion, following Merleau-Ponty, that "depth should really be called the 'first

[26] David Morris, *The Sense of Space* (Albany: State University of New York Press, 2004), 125.

dimension' rather than the 'third'; that is, depth is the most primordial dimension, not a 'bonus' dimension added to the other two."[27] Drawing on this claim, my central concern in this study is not what space is ontologically, but how we perceive spaces to be—what qualities they possess for us at any given moment.

Morris explains how we draw upon the range of habitual actions performed by the body, and our own knowledge of what our body can do, in perceiving the volume of space: "The size of my room is encapsulated in the envelope of movements I use to traverse it.… A dancer… sizes up a new stage in terms of strides and movements; a driver… maintains following distance not in terms of meters and so on, but in terms of a space of movement between cars."[28] Such knowledge is what Merleau-Ponty refers to as the body's "I can," a knowledge entailed in the very condition of having a body (in fact, he considers this knowledge of bodily capability as consciousness itself.)[29]

Musical sounds can seem to occupy various volumes of space which affect our sense of depth. In addition, there are works that confound our ability to locate sound sources, which are illustrative of the limited applicability of the space-as-container concept when it comes to music. For example, standing waves that create acoustic focal points throughout a listening area, such as those deployed in Éliane Radigue's music, make space available to our hearing. A three-dimensional "container model" of space cannot explain what happens when the sound source can only be located within the body, or within the ear itself. Such works require us to rethink how we listen and how we interpret the music that we hear.

27 Morris, 1.
28 David Morris, *The Sense of Space* (Albany: State University of New York Press, 2004), 125–26.
29 Merleau-Ponty, *Phenomenology of Perception*, 2012, 159.

1.2.5 Space and Spatial Music

Spatial music offers a good beginning point to explore how music affects our sense of space. It allows us to consider the contributions of the recorded medium to our spatial perceptions, as well as how the medium's characteristics can be put to aesthetic use. The history of the style is well documented, as is its development alongside the techniques of electroacoustic music. Spatial music is generally considered to arise with the development of recording technology and the introduction of stereophony.[30] This, in turn, opened up a new frontier for composers to explore: the orientation and movement of musical sounds. These pioneering composer-explorers were electroacoustic musicians, creating fixed-media works on magnetic tape, and not infrequently manipulating the playback signal during the work's performance. The electronic innovators of WDR and IRCAM—such as Stockhausen, Pierre Boulez, Karel Goeyvaerts, and Gottfried Michael Koenig—were not the first composers to make use of sound location as a structural feature of their works, nor was the use of location as a manipulable parameter limited to electronically processed music. Accounts of the history of spatial music typically include a token mention of Renaissance antiphonal music as an early precursor. But such accounts do not recognize this antiphony as having any influence on the extensive electronic manipulation of sound sources on display in the music of the mid-20th century. As Patrick Valiquet has persuasively argued, mid-century avant-garde composers had a vested interest in claiming that the domain of spatial manipulation of sound sources was exclusively theirs: "In order

30 Though the word "stereo" is commonly used to refer to a two-channel setup, the term in fact encompasses all multi-channel systems. The term binaural is generally reserved for distinguishing a specifically two-channel recording system. Greg Milner, *Perfecting Sound Forever: An Aural History of Recorded Music* (New York: Faber and Faber, 2009), 143.

to capture and consolidate new institutional ground, aesthetic boundaries needed to be erected on two fronts, against mass cultural banality on the one hand, and against bourgeois art music tradition on the other."[31]

Stockhausen made a point of expounding at length on the new technological resources, hailing the unprecedented character of the music they made possible, and predicting they would have a featured role in the music of the future.[32] Indeed, Stockhausen's influence is felt to this day, as spatial music is still understood primarily as music in which the location and movement of instrumental sounds is a prominent characteristic of the work.[33,]

But these compositional resources were not new. Prior to the development of sound recording technology, we can find numerous musical examples that make use of voices and instruments in separate locations. Richard Zvonar, who traces the practice back to antiquity, recounts its continued, if infrequent, use into the Middle Ages and through the

31 Valiquet, "The Spatialisation of Stereophony."

32 "His [Stockhausen's] 1958 lecture *Musik im Raum* ... opened with unqualified dismissals of every imaginable precedent for spatialisation in the western classical canon.... Gabrieli he decries as repetitive, Mozart as banal, Berlioz as melodramatic.... Stockhausen spatialised his own work in part by 'de-spatialising' that of his forebears and contemporaries." Valiquet, 414.

33 For example, Paul Miller (2012) defines spatial music as "the movement of sound in the space around listeners," and writes that it "operates in a space which has a finite or infinite set of points ("locations") in a one-, two-, or three-dimensional field;" in the New Grove Encyclopedia, the entry on Henry Brandt defines spatial music as "music for spatially separated groups." Paul Miller, "An Adventure into Outer Space: Stockhausen's *Lichter—Wasser* and the Analysis of Spatialized Music," *Perspectives of New Music* 50, no. 1–2 (2012): 342, 379. "Brant, Henry | Grove Music," accessed April 11, 2018, http://www.oxfordmusiconline.com/. The style of music that is sometimes called "space music" (as opposed to "spatial music") is a distinct, though not entirely unrelated, popular genre. Here, the word space is used in both the sense of outer space and inner space. Victor Szabo has explained how the term and its associated repertoire developed over time and that its ambiguity was purposefully cultivated by the genre's principal promulgator, Stephen Hill. As a tool for personal enhancement, the music was intended to help the listener achieve a meditative, inward state of mind, while also fostering the type of consciousness-raising experience often associated with extra-terrestrial travel and trans-personal visions of the cosmos. (Victor Szabo, "Tuning into the New Auditory Consciousness: *Music from the Hearts of Space*'s Ambient Archive 1973–83," paper presented at the Annual Meeting of the American Musicological Society, Rochester, NY, November 2017.)

Common-Practice era. According to him, antiphonal hymnody has existed since biblical times, and was used in the Christian church since the fourth century.[34] The earliest notated works that specify separation of different members of an ensemble can be found in the 16th century polychoral practices of Willaert and Gabrieli at St. Mark's cathedral in Venice. But because such music was composed for performance in highly reverberant basilicas and churches, it is unclear to what degree the separation of sound sources was discernible by congregants.[35]

Distance effects, achieved by locating musicians offstage, were employed throughout the 19th century. Beethoven's *Leonore Overture no. 3*, op. 72b (1815) features an offstage trumpet solo. Both Berlioz's *Requiem* (1837) and Verdi's *Requiem* (1874) use offstage brass ensembles. Mahler places a number of instruments offstage in his Second Symphony (1888-94), a technique extended into the 20th century by Charles Ives in his 1908 work *The Unanswered Question*. Acoustic works continued to employ nontraditional performance arrangements of audience and musicians throughout the 20th century, placing musicians in a variety of locations about the concert hall as, for example, in works by Henry Brandt, Anthony Braxton, and Iannis Xenakis.

Stockhausen famously exploited the spatial resources of three orchestras surrounding the audience in *Gruppen* (1955-57), most notably with the brass chords that are passed from one orchestra to the next. But in contrast to the works mentioned above

34 Richard Zvonar, "A History of Spatial Music," *eContact* 7, no. 4 (2004), https://econtact.ca/7_4/zvonar_spatialmusic.html.

35 David Bryant has argued that spatial separation was not a formal requirement of the majority of works performed in St. Mark's, and on the occasions when it did occur, it was small and of little acoustic consequence. David Bryant, "The 'Cori Spezzati' of St Mark's: Myth and Reality," *Early Music History* 1 (1981), 165–86.

that play with the concept of distance, *Gruppen*—like *Gesang*— attempts to give movement to sounds, that is, to make them appear to as if they were changing position in a linear manner from one location to another. This spatial feature seems to be a direct result of Stockhausen's experience with electronic studio effects associated with recording technology and of the manipulation of sound location and trajectory. Several decades after radios had become ubiquitous in homes throughout the Western world, and at the peak of the hi-fi craze in the United States, *Gruppen* demonstrates that, by mid-century, both acoustic and electro-acoustic music employed an identical conception of space. The fact that *Gruppen*'s spatial structure reflects back so effortlessly on spatiality in recorded music demonstrates the dominance of this understanding of "spatial" in Western music: a space that is a container, of variable size, furnished with sound-emitting objects at identifiable locations relative to the listener.

1.2.6 A Note on Space and Place

When describing qualities of sound that give rise to the sense of space, we often refer to the acoustic properties of particular types of places, for example, the reverberation of a concert hall, or an echo in a large church. Indeed, the notion of place is often bound up with space. Edward Casey considers place to be the primordial experience from which space is derived as a secondary, analytic or reflective quality.[36] He shares this consideration of place with other philosophers of space: Henri Lefebvre and Gaston Bachelard.[37] Yet there

36 Casey, *The Fate of Place*, 1997.

37 Henri Lefebvre, *The Production of Space* (Cambridge, Mass., USA: Blackwell, 1991); Gaston Bachelard, *The Poetics of Space* (New York: Orion Press, 1964).

is no easy way to define either space or place. In both philosophy and audio engineering, the term "place" is generally used in the sense of a specific site or location. In the present study, I will be dealing primarily with impressions of space, but oftentimes in music spatial characteristics are deeply imbricated with notions of specific locales—a home, a church, a studio—and thus make place a difficult phenomenological category to ignore.

Casey is not able to provide simple definitions of these terms either, but instead gives an exegesis of what influential Western thinkers have written about them. As mentioned above, Plato thought of space as a receptacle. Aristotle, on the other hand, considered space as something that enveloped an object. Yet there was more than one way in which Aristotle considered an object to be *in* space, and it is here, already, that some of the difficulties start to arise with attempts to define the nature of space. Casey tells us that Aristotle specified eight different modes of being *in* something: "Two of these can be considered logical or classificatory, two are metaphysical, one is political, two delineate part-whole relations, and a final one is expressly descriptive."[38]

The Platonic and Aristotelian conceptions of place have a significant posteriority in contemporary thinking on the subject. Geometry provides a model for several early modern notions of space that are even today, in the twentieth century, pervasively operative at the level of common sense, if not of scientific thinking. And the Aristotelian alternative is the active ancestor of those phenomenological approaches that, in the writings of Husserl and Merleau-Ponty, question the superimposition of geometry and call for a recognition instead of the world's immanent shapeful order.[39]

38 Edward S. Casey, *The Fate of Place : A Philosophical History* (Berkeley: University of California Press, 1997), 54.
39 Casey, 57.

Through a delineation along methodological lines, Casey ascribes to the word "space" the abstract, objective, Cartesian coordinate system familiar to physical science, while the term "place" is reserved for the family of phenomenal, perceptive, and affective relationships with one's surroundings. This does not mean, however, that space cannot be approached from a similarly phenomenological perspective. David Morris does just that in his book *The Sense of Space*, and it is this latter approach that I will follow here.[40]

As mentioned above, the phenomenological approach considers the perception of our surroundings to be inherently affective. Though the present study is concerned with the perception of space afforded by music, in the analyses that follow, I will occasionally appeal to the notion of place as well. Generally speaking, the term "place" will be reserved for the description of locales: home, studio, church, plaza, etc. While I will take care not to confound the meanings of the two, it must be understood that neither is fully extricable from the other. And, in fact, this relationship itself will be explored in my analyses, where I will show how even the musical concepts of space that are grounded in the place-related characteristics of location and direction can bring forth awareness of the affectively accessed idea of space that will be detailed in Part II.

1.3 Framing Space

The concept of "framing," as elaborated by Mark Hansen, provides a useful tool for identifying how music can contribute to the formation of our spatial impressions. Hansen's focus is on digital media and virtual reality, or VR, systems in particular. Framing refers to a

40 Morris posits an objectivized space as the abstracted counterpart to an inhabited space. Morris, *The Sense of Space*, 2004.

pair of strategies by which images are made perceptible to a spectator. One is the boundary constructed around the work to that separates the aesthetic object or activity from the ordinary, everyday world surrounding it. The other usage of framing is as a contextualizing strategy used by spectators to make sense of their surroundings and the things the encounter within them.

Hansen's theory is grounded in a phenomenology of perception that reveals the role of the spectator's body in the framing process of the work of art. His studies on electronic media apply well to the listening situations considered here, which involve fixed media and recorded works. A frame marks out the boundary of an artwork—a concept not limited to the picture frame. For temporal works such as musical compositions, a conductor's gestures can signal beginning and endings, framing works in opposition to the sounds of shuffling and tuning that precede tham and the extra-compositional applause that follows. Frames are used in interpersonal interactions as well. For example, in everyday speech, a variety of cues, linguistic and paralinguistic, are deployed to frame humorous verbal sparring as non-serious play—in other words, to distinguish a joke from a threat (such framing would also include subtleties that indicate a humorous remark might better be considered a veiled threat disguised as a joke). We can easily understand this process at work in a performance of John Cage's *4'33"*, where a variety of framing mechanisms (performance in a hall, on a stage, complete with a featured soloist and the piece and performer listed in the concert program, along with the performer's deliberate gestures—at the premiere, at least—involving the piano lid) conspire to transform the ambient sounds of the hall into an aleatoric musical work. These framing devices are necessary for a

performance to occur, but not part of the musical content. Yet this limit case exemplifies just how important the framing of the listening event is to the perception of the music.

Regarding the framing of performances, a key text is Richard Schechner's *Performance Theory* (1988), developed out of his efforts to combine theater studies with the anthropological study of ritual. Schechner's claim is that performance is an activity governed by a set of rules, rules that serve to set the performance apart from everyday life. Schechner provides the following example: "When people 'go to the theater,' they are acknowledging that theater takes place at special times in special places." The performance event is further framed by rituals such as ticket taking and finding a seat, and even extends to having dinner before or drinks after. What is notable for the purposes of music analysis, is that by treating performance as a structured activity, Schechner's theory focuses on descriptions of context and of participant experiences.

Still, exhaustively defining context can be a never-ending task: Richard Bauman and Charles Briggs call it a process of "infinite regress." But they note how work by Erving Goffman in sociology and John Gumperz in anthropology ameliorate this problem by focusing not so much on context itself as on the processes involved in constructing and participating in the context. Moreover, because these contexts are enacted performatively—that is, constructed "in-the-moment" by interlocutors acting in concert—they essentially turn any social interaction into a performance. This broader sense of performance extends beyond spoken discourse: it applies to any situation in which there is a text to be interrogated, or more to the point, an expression that warrants listening to. Any encounter that includes a "framed" musical work is similarly marked as a performance of sorts—as a contextualized practice that has its own conventions for interpretation.

Therefore, we might consider a symphony played by an orchestra at a concert hall to be framed as a "'classical' music concert" that would entail myriad social rules governing not only concert-goers' behavior, but also what they are to listen for—namely, the sequence of sounds emitted by the musicians onstage between the maestro's first downbeat and the lowering of his arms to his sides after the musicians have ceased playing. We would also consider a performance of *4'33"* as a similarly framed work, though one which directs the audience's attention to a different set of sound components that comprise the work. While the performer onstage may not provide sonic material for the audience to hear, the framing of the event as a music performance solicits from audience members the activity of listening for something, and in pursuit of that activity may well settle on ambient noises occurring in other parts of the concert hall. Thus, an ostensibly silent composition (judging from the score) is revealed to be a hushed character piece.

Hansen takes Goffman's concept of framing beyond the study of language and society, and applies it toward the cognitive domain of individual perception. In the process, Hansen draws on Maurice Merleau-Ponty's phenomenological theory of perception. As noted briefly above, for Merleau-Ponty perception is never a matter of the absorption of data by a passive observer but is necessarily a process of bodily engagement by an active subject. Hansen builds on Merleau-Ponty's work and shows how the body plays a role in framing the information available to it, a role he sees as being thrust upon the body due to the amorphous form of digital data (consisting essentially of complex series of 0s and 1s). His contention is that because a stream of digits can be an image of anything—a picture, a movie, a novel, music—there is an image only when a human subject intervenes, making sense of the data through bodily gestures. An analog example might be of use here. Artist

Olafur Eliasson's 1993 installation *Beauty* consists of a darkened room that holds a spotlight that shines on a fine mist of water falling from a device on the ceiling. As the water refracts the light, rainbow colors are visible to the viewer. These rainbows may be more or less distinct, more or less complete, depending upon where the viewer is standing. If a viewer catches a glimpse of a rainbow, and desires to see more of it, or to see it more clearly, or brighter, or observe its shimmer as air currents in the room generated by spectators cause eddies in the mist, that viewer will move their body accordingly. It is the viewer's affective concerns that motivate movement of the body—gestures—and through those gestures images emerge.

As this example illustrates, the bodily framing of phenomena is not limited to the digital realm. It is a role the body is accustomed to playing, wherever it needs to make sense of its surroundings. As Merleau-Ponty explains: "The gesture of expression delineates what it intends, makes it appear 'outside.'" The body, in marking something "out," thereby provides the frame necessary for the image to emerge. The contextualizing function of a frame, along with the body's role in its construction, will prove to be a powerful tool to help us identify what it is that is spatial about a musical work, and how it is that musical recordings can affect a listener's sense of space. By providing a role for the perceiver in the framing of phenomena, and not delegating the frame's contextualizing function entirely to the work's mode of presentation, Hansen's theory allows us to explore the physical interaction through which the work emerges. As noted above, the notion of frame will be used to describe both how our own potential for movement is called upon in the activity of listening, and in particular, how we are situated when listening to reqcorded music.

Another key term Hansen uses that is relevant to depictions of space in audio recordings is representationalism. With respect to media technology, Hansen writes that representationalism is a view that sees technology as leading to higher and higher levels of representational simulation, a view that advances verisimilitude as the ultimate aim of technological achievement. In contrast to this view, Hansen proposes that advances in technology open up the body to a wider range of possible interactions with the world, as defined with respect to the body's "perceptuomotorism," and thus to enhanced possibilities for human agency. Rather than a simulation indistinguishable from reality, technological advances in media "underwrite a more expansive and fluid...interpenetration of physical and virtual spaces.[41] When Hansen speaks of virtual space, we can understand that term as referring not only to the visual worlds created by VR systems, but also to aural ones created by recordings.

Rather than speak of "virtual" reality when considering such media simulations, Hansen suggests the term "mixed" reality, an environment composed of several channels of information, each conveying an image of reality, and each separately accessible. For example, players of immersive video games still understand themselves as playing a game. What technical sophistication can bring to the experience is a *seamlessness* that characterizes the manner in which the reality of the game is woven into the game player's everyday reality, an interpenetration of virtual space and actual space.[42] This separation and access is in accord with Schechner's performance theory. According to Schechner, however immersed in a performance one might be, one never loses one's perspective as

41 Mark B. N. Hansen, *Bodies in Code: Interfaces with Digital Media* (London: Routledge, 2005), 3.
42 Mark B. N. Hansen, *Bodies in Code: Interfaces with Digital Media* (London: Routledge, 2005), 2.

audience member. That is, one never actually believes that the actor playing Oedipus has actually gouged out his eyes, rather than just pretending to have done so. Similarly, however immersed in music one might be, it would seem that the "spectator" role of a listener must be maintained for there to be a reflective interpretation of the experience. This distance is a result of the distinction maintained between the "here-and-now" of the spectator and the "there-and-then" of the actions on stage, and which thereby allow for a reimagining of the theater-goer's world in light of the plot of the play.

Hansen's proposition that technology underwrites the interpenetration of separate streams of information is a useful lens through which to view our interactions with music because it allows reflective judgment to enter our experience. Where we are situated—what the context is—is an affective judgement. Therefore, what we perceive, what we pay attention to, what we frame as an object of interest, all these involve an affective component. As our needs and desires change, so does our perspective. Thus, music can provide opportunities to experience our subjecthood differently by affording different experiences of our spatial environment.

1.4 Historical Recording Practices and the Presentation of Space

Just as framing conventions affect what we hear as music, so does the technology employed to present it. Recording technology is associated with a set of listening practices that are the result of a particular history of its development, marketing and use. The structures we hear in music are determined not just by the content of the music, or even a combination of musical content and the constraints of the medium, but also by what it is that we are listening for. Because how we listen and what we are drawn to listen to are so important to the sense we make of the sounds we hear, we need to understand what some of these practices are, how they developed. Once established, such practices achieve a naturalness that often lets their contingent nature pass unnoticed. They might therefore obscure some of these spatial features we are trying to identify.

There are several reasons for focusing on recorded music in a study of space in music. One is that much theoretical work depends upon multiple listenings of works, and upon repeated close listening to individual passages. While this poses less of a problem for piano music, or indeed any solo work analyzed by a theorist sufficiently proficient on the relevant instrument, repeated listening is not feasible for orchestral works. Another reason is that much of what is termed "spatial music" consists of fixed media works, recorded onto magnetic tape or computer disc and played back through loudspeakers. Third, it is evident that recording practices have had an influence on analog musicking practices. In addition to the example of "Gruppen" above, that shows the influence of studio panning techniques, witness the abandonment of traditional dynamic markings "*crescendo*" and "*diminuendo*" in favor of the studio terminology "fade in" and "fade out" in the music of Steve Reich. A fourth

reason is that spatial considerations were highlighted by the recording industry from early on in its history: debates over recording techniques were often centered on how to handle the acoustics of the recording environment. This resulted in the development of electronic processing units that could simulate the acoustics of various interior spaces, with important consequences for how that music was heard by the listener. Finally, phonographic equipment was marketed as a consumer product designed for home use. (It could have been otherwise: Phillip Auslander notes a scheme in the late 1940s for installing TVs in movie theaters for public viewing by paid admission.) Playing music in the home instantiates an interpenetration of spaces: that of the recording and that of the listening area. This unique situation gives rise to experiences that highlight music's role in creating space as an audible, sensible phenomenon, as well as shedding light on how we interpret the music we hear.

The historical contingency of recording practices is examined in Jonathan Sterne's *The Audible Past* (2003).[43] Sterne identifies a number of assumptions inherited from antiquity and the Middle Ages regarding the relationship between sight and sound and groups them together in what he calls the "audio-visual litany," for example the idea that "hearing is a primarily temporal sense, vision is primarily spatial," or that "sounds come to us, but vision travels to its object."[44] Simply by tracing the historical origins of these preconceptions, he begins to undermine their fixity and naturalness, thereby making his broader point that categorical assumptions about perception and technology are open to critique. Sterne is particularly effective in explaining one of the first uses of sound

43 Jonathan Sterne, *The Audible Past: Cultural Origins of Sound Reproduction* (Durham: Duke University Press, 2003).
44 Sterne, 15.

recording: as a permanent record of the deceased. He ascribes this to the fact the technology was developed in the midst of a "late-Victorian death culture in the United States and the United Kingdom " that sought and sanctioned just such a use for it.[45] Citing the practices of canning and embalming that had come into widespread use by the latter-half of the 19th century, Sterne writes, "hopes that phonography would preserve the voices of the dead were only an extension of a larger, emergent culture of preservation."[46] Though Thomas Edison eventually turned his attention toward recording music, Sterne relays how Edison's original vision for the phonograph's use was as a family record: "a registry of sayings, reminiscences, etc., by members of a family in their own voices, and the last words of dying persons."[47] (It would not be until a later date that competition and economic concerns prompted the practice of recording music.)

Sterne also describes the construction of a particular conception of sound recording—sound reproduction as a relation of equivalence between an original and a copy. Public demonstrations by Edison Phonograph Company in the late 1910s and early 1920s, pitting singers against cylinders, were theatrical confections based on commercial schemes that marketed the early product as fulfilling this new equivalence function, a function that had become the *sine qua non* of recording. Conceiving of the device in this manner then led to technological developments that were able to realize this aesthetic goal. Greg Milner elaborates on Stern's account of these demonstrations and introduces a key

45 Sterne, 288.

46 "Modern canning, which began early in the nineteenth century, did not become widespread until shortly before and during the American Civil War. The mass production of tin cans, beginning in 1849, Borden's method for canning milk patented in 1856, and the inventions of screw-top mason jars and bell jars in 1858 helped stimulate the spread of artisanally and industrially produced canned goods." Sterne, 292.

47 Sterne, 305.

concept in recording that had an impact on later stereophonic recording practices: verisimilitude. Milner's history highlights the budding industry's concern with realism, veracity, and purity with respect to both musical works and musical performances. Such concerns helped define the direction of technological innovation for many decades, though just what features of the recordings served as markers for the real, the true, or the pure, varied over time and among artists and engineers.[48]

Like Sterne, Milner notes that Edison intended early on for his invention to be used to preserve sounds, believing that music might be among those sounds that would benefit from preservation. (However, Edison envisioned for the phonograph being used primarily for the preservation of speech. Music ranked fourth out of a field of ten: the other uses were all based on spoken word recordings.}[49] But Edison also considered the phonograph to be a "truth-teller," a machine with perfect recollection, such that it could be used, as he suggested, for business conducted on the telephone, obviating the use of written contracts.[50] Similarly, he maintained "it would even be worthwhile to compel witnesses in court to speak directly into the phonograph, in order to thus obtain an unimpeachable record of their testimony."[51] Later on, in the 1910s, when recording and playing music became the primary commercial use of phonography, Edison continued to extol the

48 Milner and Sterne both believe that the potential of the medium is not inherent in the technology itself, but that it lies in the uses to which it is put, which in turn guides the technology's future development. Neither author considers what media theorist Gilbert Simondon calls the "autonomy" or "internal coherence" of technological inventions as a factor which generates innovation. Pascal Chabot explains that for Simondon: "Once in operation, the technical object frees itself from its inventor. Its superabundant functionality separates it from any plans or intentions projected on to it." Pascal Chabot, *The Philosophy of Simondon: Between Technology and Individuation* (London: Bloomsbury Academic, 2013), 15.

49 Thomas A. Edison, "The Phonograph and Its Future," *The North American Review* 126, no. 262 (1878): 531–535.

50 Edison, 535.

51 Edison, 533.

phonograph's truth-telling abilities. More than just providing listeners with a practical equivalent of a musical performance—recording as documentation—Edison claimed his recordings had the ability to present the *truth* of a musical work. .[52] Thus, along with the preservation motif identified by Sterne, there arose a trope of veracity, a trope which—through Edison's "tone tests"—became intertwined with portrayals of realism.

"Tone tests" was the name Edison contrived for his phonograph demonstrations conducted from 1915–1925. Milner describes how these events were cast not as concerts, but as something like magic shows. Audience members were present at a performance, but not necessarily, or at least not entirely, a musical one. They were prepared to witness acts of virtuosity, whether by musicians or machines. The sense of awe these performances were designed to induce in audience members set the stage for future listening practices. The tone tests featured professional musicians matching their talents against Edison's latest phonograph model. But, as Milner describes it, the star musicians brought in for these occasions were not the main focus of the event: "the real headliner was already onstage, front and center: an Edison Diamond Disc Phonograph in a wood cabinet."[53] (One can imagine a scene not unlike that depicted in Figure 1: an expectant audience with eyes riveted on a box sitting on a stage.) Demonstrations would consist of a singer starting a song beside the phonograph. After suitable interval, the singer would close their mouth, but the singing would continue—with gasps from the audience when they could see that the performer had stopped, but that the music they were hearing did not. These tests, however, were rigged: Edison instructed the singers to match the tone of their voice to the tone of

52 Milner, *Perfecting Sound Forever*, 35–36.
53 Milner, 4.

the machine—mimicking as best they could the constricted, tinny color of the sound coming out of the horn. Sterne points out the semiotics of staging at these events:

> Here, for the first time, was a live performer and a sound-reproduction device presented onstage—together as equals. The metonymic logic was clear enough—if these great performers can share a stage with the Edison phonograph, then live musical performance and recording can be understood as two species of the same practice.[54]

While the equivalence relation Sterne refers to here is better understood to be one between original and copy, his reading is still an accurate one: these demonstrations were teaching audiences how to listen. Sterne calls this *audile technique*: "a particular kind of listening for detail and a particular relation between listener and instrument."[55]

These demonstrations serve as an example of how a particular listening practice emerged through the framing of the technological demonstration as a music performance, one that ignored the limitations of the medium (pushing them aside, or out of the frame) while reveling in its novel capabilities. Benjamin Steege, in his account of the cultural beliefs and social values underlying 19th century experimental science, shows how aesthetic perception had been "discovered and manufactured, altered and prolonged, regimented and liberated."[56] In describing the development of Hermann von Helmholtz's psychophysiological proposition that our hearing abilities are both malleable and educable, Steege shows how such a historically contingent viewpoint is constructed, and

54 Sterne, *The Audible Past*, 262–263.

55 Sterne, 223. Stockhausen also made use of staging to prepare his audience for the premiere *Gesang der Jünglinge*. As with the Edison tone tests, Figure 1 shows that the technology itself was on display: a behemoth of wood and wire poised to deliver a virtuosic solo performance. But rather than forging an equivalence relation between recording and live performance, Stockhausen's setup removed the distinction altogether, while strictly enforcing codified listening practices.

56 Benjamin Steege, *Helmholtz and the Modern Listener* (New York: Cambridge University Press, 2012), 15.

consequently, the malleability and educability of our conception of hearing itself.[57] He also reminds us that it was Helmholtz himself who revealed the apparent selectiveness of hearing, its "forgetting" of detail (specifically, a tone's overtones) and how what is heard is often determined through habituation rather than through attention or reflection.[58] Helmholtz's use of a resonator that could highlight component frequencies of a tone while muting others shows how the resources of an institutional apparatus could be employed to reshape a listener's hearing. Through the use of resonators—employed within a scientistic context—Helmholtz's listeners were trained to hear differently, to interpret what they previously considered a single tone to instead be a compound sound made up of a combination of several different frequencies.

Just as Helmholtz was able to inaugurate a novel understanding of tones and to redescribe listeners' experiences of what they heard, so were the manufacturers and vendors of phonographs able to shape phonograph listeners' understanding of what they were listening to. Promotional materials provided instructions to consumers about how their product was intended to be used. For early recordings, it was necessary to prepare listeners for what they were about to hear.

Mark Katz describes the special conditioning required for a recording to be perceived as realistic. He depicts the elaborate ritual surrounding in-store demonstrations of Edison's phonograph by local distributors:

> "Picture the scene," you are told, until "it is clearly . . . in mind." Once this mental image is firmly in place, you are to say, "I am ready," at which point the demonstrator plays your chosen record. The final instruction is wonderfully complicated: "About forty-five seconds after the music begins, close your eyes and keep them closed for a minute or more. Then open your

57 Steege, 44.
58 Steege, 73–75.

> eyes for fifteen seconds but do not gaze at your surroundings. After this, close your eyes again and keep them closed until the end of the selection." If you follow these directions exactly, you will supposedly get "the same emotional reaction experienced when you last heard the same kind of voice or instrument." If for some reason you do not, it is because "you have not wholly shaken off the influence of your surroundings."[59]

The responsibility for a proper hearing is placed squarely upon the listener's shoulders; as the sales pitch makes clear, the technology is faultless.

It was not only the listener's surroundings, however, that Edison considered a distraction; with respect to music, the real word itself was an encumbrance. He believed that what recording ought to document is the sound of music in a non-reverberating utopia—a "no-place"—that would not admit any noise he considered extra-musical to come between the music and the listener.[60] For Edison, this musical truth could best be delivered through recordings that de-emphasized ordinary room acoustics, that is, recordings that contained a minimum of reverberant sound.

But this was not the only means to attain a recording that was purportedly true to the music. Edison's goal of purifying musical sound was in contrast to that of his competitor, the Victor Talking Machine Company, manufacturer of the popular Victrola phonograph. The Victor Company took account of the limitations of the hardware, and engineered recordings to have a resonance that was considered to enrich the tone and appeal of the music. Milner, quoting phonograph collector Michael Devecka, describes the effect:

> Victor tried to record the overtones, the noise of the room. The idea was to give it a little more warmth, a little more resonance, kind of the singing-in-

59 Mark Katz, *Capturing Sound: How Technology Has Changed Music* (Berkeley: University of California Press, 2010), 26.
60 Milner, *Perfecting Sound Forever*, 139.

> the-shower effect. Edison, on the other hand, wanted to record in a deader studio. He wanted accuracy. He didn't want a lot of extra stuff going on.[61]

The Victor Company understood room acoustics to be a valuable component of the overall sound of recorded music; not for the sake of realism, however, but as a positive timbral contribution, an added "warmth." To Edison, these same acoustic qualities got in the way of the music, covering it up by making it less focused and precise.[62] These opposing paths toward the same goal demonstrate that there is no single set of features that can be used to characterize the faithfulness of a recording or its resemblance to an ideal. More importantly, the spatial qualities of music can never be considered independently of the choices made in the recording, nor of the social and cultural conventions we rely upon in order to approach it.

1.5 A Domestic Technology

As Sterne noted, music was not the first purpose Edison envisioned for his new technology. It took a couple of decades for manufacturers to come up with a profitable commercial use for their products. After he filed his patent in 1877, Edison's first attempts to market the phonograph were unsuccessful.[63] But after the development of more durable wax cylinders, the North American Phonograph Company decided to capitalize on Edison's stated "main utility" of phonography: dictation. Mimicking the successful business model of

61 Milner, 57.

62 It might be said that even for Edison, realism was a function of room acoustics—it was simply that for Edison's goals, the room was a smaller one, with surfaces that tended to absorb sound rather than reflect it. From his statements, however, Edison did not likely see it this way.

63 Leonard DeGraaf, "Confronting the Mass Market: Thomas Edison and the Entertainment Phonograph," *Business and Economic History* 24, no. 1 (1995): 89.

the telephone industry, the company initially marketed its product as "a tool for business communication, designed to replace stenographers."[64] After this failed to prove viable, Edison's National Phonograph Company reimagined the device as a public novelty, to be listened to through headphones at penny arcades. With the phonograph still conceived of as a talking machine, not least due to the limited frequency spectrum it could capture, many of the cylinders made for public phonographs were spoken-word recordings—humorous stories and dramatic monologs—along with recordings of folksongs and siffleurs (professional whistlers).[65] Musical selections were typically solo pieces for instruments that happened to record well, such as the piccolo and cornet, as well as wind band music, since brass instruments recorded relatively well too.

Though a successful money-making endeavor for a few years, the arcade device eventually lost its novelty. In an effort to boost sales, in the first decade of the 1900s, the recording industry mounted a concerted campaign to establish the phonograph as a domestic device, in large part by designing it as furniture, so that it could be accommodated into the middle-class household. With the ability to mass produce recordings developed by Emile Berliner in 1895, and with the phonograph staking out domestic terrain as its milieu, new relationships to music were opened up for listeners. As Katz explains, phonographs let their owners hear music that they might otherwise never have the opportunity to hear, allowing "'good music' to be brought into the home."[66]

64 Kyle S. Barnett, "Furniture Music: The Phonograph as Furniture, 1900–1930," *Journal of Popular Music Studies* 18, no. 3 (December 2006): 304, https://doi.org/10.1111/j.1533-1598.2006.00096.x.
65 James R. Smart, "Emile Berliner and Nineteenth-Century Disc Recordings," *The Quarterly Journal of the Library of Congress* 37, no. 3/4 (1980): 430.
66 Katz, *Capturing Sound*, 57.

"Bringing music into the home" set phonograph use beside other activities like bringing in guests or bringing in groceries, or more fancifully: "it was believed that in bringing classical music into the home, the phonograph also brought refinement and taste."[67] Likewise educators saw phonographs as a less expensive choice between two alternatives: "[Music] Appreciation ... required that either music be imported into the classroom or the children be transported to the concert hall."[68]

Thus, a concept of both music and listeners as objects that can be moved from one location to another took hold and informed listening experience by providing an explanation of the purpose of the technology. These functions assigned to the phonograph—of "importing" music and "transporting" children—established recorded music as something more than simply a commodity. It was portrayed as a service as well. These functions fostered the belief that recorded music was a representation and recreation of performance, whose essential features are precisely those that are captured by the recording. Though in one case music may be performed live, and in another it may be played back from a wax cylinder, discourse such as that cited by Katz reinforced the belief that it is the same music either way.

67 Katz, 57.
68 Katz, 42.

1.6 Verisimilitude: The Living Room vs. The Concert Hall

The theme of transport and distribution that Katz identified persisted long into the twentieth century. Once established in the home, phonographs were touted as being able to bring sounds of the concert hall into the living room, and conversely, to transport the listener into the concert hall. Greg Milner calls this the "living room vs. concert hall dialectic": competing concepts of what counted as verisimilitude in recorded music.[69] Should listeners feel that the performers are in the same room as they are, or should they have the experience of being seated in a concert hall? Kier Keightley notes how coverage of the "hi-fi craze" of the 1950s documented a keen interest in the former scenario among home listeners:

> The "typical" purchaser of such records was called a "sound-for-its-sound-sake (sic)" enthusiast by a *HiFi & Music Review* article in 1958 (Jacobs 1958, p. 33). The article begins: "Addict or realist? This is the question neighbors ask when assaulted by 'weird' and inexplicable noises from that hi-fi system next door. Some hi-fiers, rather than immerse themselves in operatic or chamber music, or even rock 'n' roll, listen for the joy of just 'hearing' sounds not likely to be found in the average living room" (Jacobs 1958, p. 33). Notice the reference to "immersion", a common trope used to describe both LP and hi-fi listening, as well as the insinuation that the hi-fi is being played at high volume.[70]

The weird sounds mentioned by the magazine refer to a predilection among hi-fi owners to judge the quality of their systems by playing records featuring nothing but sound effects. By pointing out that hi-fi's audio pleasures come as much from the magical effect of hearing unconventional sounds within the home as they do from listening to music, the review

69 Milner, *Perfecting Sound Forever*, 139.

70 Keir Keightley, "'Turn It Down!' She Shrieked: Gender, Domestic Space, and High Fidelity, 1948-59," *Popular Music* 15, no. 2 (1996): 152.

author shows that the equipment was still portrayed as a modern marvel. Higher wattage levels did allow consumers to achieve greater frequency response and dynamic range from their systems—thus minimizing limitations of the technology so that the musical content could shine through. However, it seems clear that the dazzling effect of a device that could transport the sound of recognizable objects into the home, at will, and at high volume, also held great appeal to hi-fi aficionados.

Keightley cites sources that show the other term in Milner's dialectic, the "concert hall" experience, is also represented in the popular press surrounding the hi-fi system:

> *Life* magazine offered a not dissimilar account of hi-fi in 1953: "Today the goal is to transport the listener to the concert hall" (Brean 1953, p. 156). A 1957 *Newsweek* letter to the editor suggested that hi-fi fans made an important distinction between "lo-fi" and the real thing, the latter defined by the illusion of a transparent transportation: "Do you want your radio or phonograph to sound like a radio or phonograph or do you want to be transported to the concert hall without leaving your living room? I'll take hi-fi" (Dare 1957, p. 19). *High Fidelity* magazine also described high fidelity as a virtual experience: "the usual definition of hi-fi, by its exponents at least, is the illusion of being in one's favorite seat in the concert hall...." (Wayne 1958, p. 45).[71]

Thus, along with the novelty of bringing foreign sounds into the living room, recordings were also portrayed as a means of escape: "the illusion of a transparent transportation." As with the distance cues noted above, this import-export motif reflects the idea of space as a container filled with musical objects: objects placed indoors within the vicinity of the listener (living room) or the listener placed within the confines of the recording environment (concert hall). These two functions of phonographic sound map onto the conflicting understandings of virtual reality described by Hansen. While the technocratic

71 Keightley, 153.

vision for VR promises to supplant one's own reality with the virtual image ("the concert hall"), Hansen's theory of mixed reality portrays a space ("the living room") with multiple streams of accessible information. It can thus account for the phenomenological experience of listening to recordings that contain embedded representations of the spaces in which they were recorded.

Rather than a fully convincing illusion, Hansen maintains that technological advances serve to "underwrite a more expansive and fluid *functional* interpenetration of physical and virtual spaces."[72] Hansen here opposes functionalism with representationalism. As I previously noted, representationalism is a view that sees advances in technology as leading to higher levels of representational simulation. Functionality, by contrast, is a purpose or use to which a simulation is put, a use that coordinates the virtual projection of the medium and the actual world of the viewer's body. Importantly, the virtual image itself is never a surrogate reality. On its own, it can never open up possibilities for action (i.e., it has no "function") because if the simulation did not admit of interpenetration with the actual, the virtual would be experienced as a delusion rather than an illusion. The two realms of the real and the virtual are entwined: it is their coexistence that marks the media experience as novel, or uncanny, or revelatory.

Hansen takes into consideration these layers of reality and representation. A more convincing simulation, he claims, is not what is achieved by innovative digital technologies and the specific "realities" they present to us. Rather, their ability to engage us is predicated upon our embodied condition and our capacity to fashion a reality for ourselves through "motor intentionality," or our own deliberate movement. It is the body that integrates these

72 Hansen, *Bodies in Code*, 2005, 3.

various realities into a seamless experience. Thus, Hansen's theory is not about the fragmentation of experience into segregated sensory inputs, but about an integration founded upon an "intact, coherent bodily schema" that is "always already" engaged in this integrative process.[73]

Phonograph listening directs our attention by means of the history of listening practices that are associated with it. The function of all the press, publicity, and product demos for the phonograph was in large part to coach listeners how to hear.[74] Phonography in these demonstrations was about more than simply the musical content—the virtuosity of the technology was always on display. Recording technology therefore was never really a transparent medium for music. Part of the manufacturers' marketing strategy was to convince consumers to value certain aspects of the product while ignoring others. The use of a recording as a permanent record of an event or as a communication medium was nudged aside to make room for the idea of the recording as a convincing recreation of a distant event. And the phonograph, rather than serving as an archival tool, was to be understood as a musical performer in its own right.

The significance of this realignment cannot be overstated. The recording apparatus as a machine that performs music, or as a type of self-playing musical instrument, is essentially how it is used today. The fact that what a stereo system (or mp3 player or computer) replays is a fixed, archival product is of secondary importance. We know, for example, that the content of a CD is fixed, and that it stores a digital record of a set of musical acts performed (or programmed) at an earlier time for us to listen to when we

73 Hansen, 153.

74 Not unlike the way Helmholtzian listeners, via the resonator, were educated in both how to listen and in what to listen for. Steege, *Helmholtz and the Modern Listener*.

want. Even when we know the performance had already occurred at another time and place, we often still listen to it as an experience in which we are co-present with the music—as if it were a mode of performance as well. What makes the listening event virtual is not the degree of verisimilitude achieved by the recording media, but rather that fact that the media's content is projected alongside us in the room—present with us in the same time and place—along with the relative ease with which we are able to switch among the different channels of information available to us.

1.7 Descriptive Recordings and the Acoustics of Places

One important use of this technology is creating impressions of directional sound. The majority of these can be achieved with studio production techniques that manipulate two variables: lateral placement of sound source in the stereo field, and its distance from the listener.[75]

According to William Moylan, there are three ways in which spatial relationships are detected by the ear:

1. as a sound source located at an angle to the listener (above, below, behind, to the left, to the right, in front, etc.),
2. as a sound source located at a distance from the listener,
3. as an impression of the type, size, and acoustic properties of the environment.[76]

75 Moylan, *Understanding and Crafting the Mix*, 2007. Though Dockwray and Moore (2010) posit a vertical axis as well, it is conceived as function of pitch: relying on a convention of considering sounds with shorter wavelengths to be high relative to those with longer ones. Moylan, however, states that "currently used audio playback formats can only accurately and consistently reproduce localization cues on the horizontal plane" (p. 24). Ruth Dockwray and Allan F. Moore, "Configuring the Sound-Box 1965–1972," *Popular Music* 29, no. 2 (2010): 181–97.

76 Moylan, *Understanding and Crafting the Mix*, 2007, 24.

Numbers 1 and 2 refer to source location along two axes: the left-right, or transverse axis, and the front-back, or sagittal axis. Number 3 is a description of what Moylan calls the host environment: "the environment in which the sound is sounding."[77] The space-as-container concept is reinforced here by the reference to object position and a definition of "environment" as the geometry of a fully enclosed room.[78] Until the advent of recording technology, music's environmental characteristics were a fixed feature of the performance space. Composers could, and did, write music for the spaces in which it was to be performed (e.g., chamber music), but in general those spaces were not compositional resources to be drawn upon at will, like keys, meters, and orchestration. The exception is when musicians were directed to be seated offstage, the physical barriers to efficient sound transmission altering the sound that reached the audience.

The representation of environmental space in recording can be traced back to the 19th century. Peter Doyle (2005) notes the production of what were called "descriptive records" in the 1890s, recordings intended to represent the sound world of a particular setting or environment, as with one such record titled "Morning on the Farm" (1896): a ninety-second recording of animal noises.[79] (Or rather, *simulated* animal noises. Doyle neglects to mention that these sounds are actually imitations offered up by a human performer, as revealed toward the end of the recording when the impressionist announces: "Now this is a cat and dog fight.")

77 Moylan, 11.

78 William Moylan, *Understanding and Crafting the Mix [Electronic Resource]: The Art of Recording* (Boston: Elsevier/Focal Press, 2007), 10–11.

79 Wood, N. R., *Morning on the Farm* [United States: E. Berliner's Gramophone, 1896], audio, retrieved from the Library of Congress, https://lccn.loc.gov/99389823 (Accessed October 19, 2017). (1:27)

A farm, obviously, is an outdoor space, but rather than indicating spatial relationships, what these sounds index is an identifiable type of *place*. They do this not by reproducing the reverberant properties of an enclosure, but through a portrayal of the sound objects that might be found there. It is a depiction of a scene, just as a visual depiction of the same scene might consist of a painting of a barn surrounded by cows and sheep.

Another example of a descriptive record is "Departure of a Troopship" (1905), a dramatization of the sounds of British soldiers leaving home for the Boer War.[80] Like "Morning on the Farm," this scene was produced entirely in the studio. And again, to convey a sense of vividness, environmental sounds were used to index a particular place and a particular time. For example, by increasing the volume, a military band is made to sound as if it is approaching the listener. Toward the end of the recording, the band becomes gradually softer, accompanied by a soft steam whistle, as if the ship had sailed off into the distance. What is notable here is that the techniques for representing certain spatial characteristics on a recording were available to engineers in the early years of the technology, long before the advent of stereophony. Such descriptive records were based on the model of a visual image, such as a painting or a photograph: the objects depicted were provided with locations relative to one another within the scene.

An aural depiction, then, can be said to have a frame just like a photograph or a painting. Here, it is provided by the beginning and ending of the phonograph record, a minute-and-a-half facsimile of what a beholder would perceive if she were in the same location at the same time as the subject position indexed by the sound image. But beyond

80 Available at https://www.youtube.com/watch?v=sdX3ckozQas. (Accessed January 30, 2018.) (2:55)

the immediate boundary-defining function of the frame, Doyle recognizes a further, contextualizing function provided by the conventions of the technology's use. The sounds of a foreign scene (along with its concomitant spatial features) are brought into the home—as *objects*, like acquisitions or collectables. Similar to Hansen's description of mixed reality, Doyle characterizes the effects of a descriptive recording as intermingled but separate information channels: "This 'sound picture' overlays the fictive spatiality of the 'picturesque farm' onto the actual space of the 'parlor,' *without challenging the integrity of either space*."[81]

This frame, then, is not conterminous with the listening space, as sounds may appear to be quite remote, originating from a distance far beyond what the boundaries of the room would allow, but the boundaries of each can be considered concentric. Both a visual scene and a recorded scene have a subject-oriented perspective, and this subject position allows both spaces to co-exist in a manner that is familiar to us from viewing pictures, illustrations, and photographs.

Listeners to descriptive records experienced an altered reality in their living rooms: a refiguring of their own space as if it were a farm (or quayside). Doyle quotes a contemporary catalog description of the "Morning on the Farm":

> "So real and exact that it requires but a slight stretch of the imagination to place one's self in that delightful position, the result of which is the drinking in of copious drafts of fresh air and numerous other pleasures attainable only on the farm."[82]

81 Peter Doyle, *Echo and Reverb: Fabricating Space in Popular Music, 1900-1960* (Middletown, Conn: Wesleyan University Press, 2005), 50. Italics added.

82 Doyle, 50. The animal impressionist similarly provides a helpful hint on how to hear when he introduces his "cat and dog fight," a descriptive introduction evidently deemed unnecessary or redundant for his other imitations.

This early disc record is not a moment preserved for posterity or an aide-memoire, but a means to experience a "virtual reality" in much the same way as we use that expression today. That is, the recording provides a simulation with enough convincing resemblance to convey an impression of being in a different place. The listener is primed for this not only by the catalog notes, which tell listeners what they are to listen for (as well as admitting that true equivalence will still require "a slight stretch"), but also by the fact that the Victrola was already being tasked with conveying experiences, rather than messages.[83]

Likewise, as Keightley's marketing materials from the hi-fi era suggest, the recreation of a performance space, such as a concert hall, a cathedral, or a small club, was made to function as a marker of fidelity, or what we might otherwise call, in a musicological context, authenticity. Anna Zayaruznaya (2017) notes how, even in the twenty-first century, reverb is often added during post-production to modern recordings of medieval motets in order to "evoke a churchy 'medieval sound.'"[84] Such processing, she suggests, is the result of "modernist preferences propagated under the banner of authenticity."[85] Reverb in this case serves the twinned purposes of verisimilitude and authenticity.

The echo of the hall makes the recording "lifelike" in the sense that the long reverberation times make the music sound as if it were recorded during an actual

83 Here the distributor's text shrewdly employs a tactic recognized half a century later by Roland Barthes. Barthes uses the term inoculation to describe how faults are made trivial through their admission. Roland Barthes, "Operation Margarine," in *Mythologies*, trans. Annette Lavers (New York: Noonday Press, 1972 [1957]): 40–42.

84 Anna Zayaruznaya, "Intelligibility Redux: Motets and the Modern Medieval Sound," *Music Theory Online* 23, no. 2 (June 1, 2017), http://mtosmt.org/issues/mto.17.23.2/mto.17.23.2.zayaruznaya.html.

85 Zayaruznaya.

performance in a (presumably) appropriate venue for such music: a stone cathedral. Also, by indexing the voluminous interior and hard, flat surfaces of a medieval cathedral, the reverb might be heard to convey a sense of what the music sounded like for its original listeners (a pretense Zayaruznaya thoroughly debunks by revealing the historical inaccuracy of such a performance venue.) These conflated values of verisimilitude and authenticity often conspire to obscure the sleight-of-hand involved in attributing uncanny stereophonic effects to a meticulous simulation of the recording environment. But once again, it is not a matter of simulation, but of simultaneity—the admixture of competing versions of surrounding space.

While Doyle was writing about a long-ago listening practice, his remarks about the depicted space being overlaid upon the parlor match up well with Hansen's theory of virtuality. For Hansen, the virtual image does not supplant reality. Rather, it adds another channel that one can attend to, accessing it in accordance with one's needs and desires through "enaction," that is, through embodied perception.[86] For an appropriately trained audience, playing descriptive records provides alternative spaces to explore within the confines of a home listening situation.

If a listener experiences the sound of a cow in her living room as uncanny, the discourse surrounding hi-fi and consumer electronics assures us that this is because the ear and the eye are feeding us incompatible streams of information. And yet, according to such texts, it is always the virtual the wins out. "Close your eyes, and you are there," the promotional copy tells us; "The listener is placed in that position." Yet Hansen tells us there is another process at work, one of framing, accomplished by an embodied subject, for

86 Mark B. N. Hansen, *New Philosophy for New Media* (Cambridge, Mass: MIT Press, 2004), 2.

whom slipping in and out of different perspectives becomes, thanks to sophisticated technology, "seamless." Our body integrates the "incompossible" through its expressive gestures, and necessarily so. The shifts in perspectives are always motivated, and are accomplished affectively, guided by moment-to-moment wants and needs. The emphasis on the technological medium at the expense of the perceiver's gestures hides the affective qualities that make space *appear* for us.

Making the ability to portray sound source location the *sine qua non* of high-fidelity recording reifies the container model of space. As a result, the study of music's spatial effects has generally focused only on the "virtual"—the distance, location, and trajectory of sounds within the "container" the recording recreates—at the expense of the transformation of the actual, felt surroundings of the listener. Large and small spaces are distinguished by measurements—whether in terms of linear units like feet and inches, or the milliseconds of reverberation times—rather than the perceptions of an inhabitant of that space. But by addressing the affective nature of perception we can begin to look at how our own, actual space is transformed by the music we hear within it.

1.8 Integrating Spaces: Miles Davis, "Maiysha"

The conventional idea of a recording portraying a set of musicians in a fixed performance location persisted for many years. But as technology became more sophisticated, and a greater degree of control over the sound of the final recording became possible, musicians and engineers began to experiment with new ways of recording music—ways of representing performances that went beyond what could be achieved in a

concert setting. (Brian Kane provides an excellent case study of this use of technology for the presentation of impossible, unreal performance configurations in his account of Les Paul's multi-track recordings from the early 1950s.)[87] But even earlier, in the 1930s and 40s, Leopold Stokowski was manipulating the gain on the control panel at key moments in order to highlight aspects of structure and texture in ways that would not be possible in the concert hall.[88] In the early 1970s, jazz musician, composer, and bandleader Miles Davis and his Columbia Studios engineer, Teo Macero, challenged the fixed performance setup in even more radical ways in their collaboration on Davis's jazz fusion albums. Influenced by a variety of artists, from Sylvester Stone to Karlheinz Stockhausen, the pair began to produce pieces that highlighted the artificiality of their performance settings and that deliberately effaced some of the most common and indispensable characteristics that present a recording as a representation of live performance.

Though mid-century avant-garde composers derided the use of multi-channel recording to provide an added layer of verisimilitude via location cues, the technique was embraced in jazz, pop, and classical music recordings, where the sonic artifacts of the locations in which they were recorded were eagerly sought out and artfully arranged.[89] The sound character of actual performance events was often privileged, with efforts concentrated on reproducing what music would sound like if it were heard while sitting in a prime seat in a concert hall or jazz club. This sonic ambience was achieved either through

87 Brian Kane, *Sound Unseen: Acousmatic Sound in Theory and Practice* (New York, NY: Oxford University Press, 2014).
88 Arved Mark Ashby, *Absolute Music, Mechanical Reproduction* (Berkeley: University of California Press, 2010), 46–50.
89 Valiquet, "The Spatialisation of Stereophony."

recording music in the desired host environment, or simulating that environment in the studio through the use of artificial reverberation and mixing techniques.

The Miles Davis album *Get Up with It* (1974) combines these two approaches to music recording—the conventional and the avant-garde—to create novel spatial experiences for the listener. Experiences, moreover, that are not predictable from the spatial cues they rely on.[90] Certain moments of the album's tracks can be brought to bear on questions concerning the contribution of studio recording techniques to musical space and how both obvious electronic effects and sonic artifacts of the recording process itself affect those spatial perceptions. While trafficking in the techniques that reflect the directional, three-dimensional conception of space outlined previously, producer Teo Macero deftly combined conventional spatial cues with indices of place to create extra-ordinary musical experiences.

The second track on the album, "Maiysha," is constructed as an A B form, with the A section consisting of a 16-bar song form repeated several times as both an introduction and—as is typical in a jazz work—a framework for soloing. It features a small ensemble with Miles Davis playing electronic organ and trumpet. The characteristics of the space rendered by the recording has a distinct profile—large and reverberant, yet with microphones placed very close to some of the instruments so that the direct sound masks any reflections off the walls of the room. These walls seem highly reflective, as evidenced during the moments when the texture thins out and the claves are heard to possess a considerable ring (for instance, at the beginning of the chorus that starts at 2:58). Whether

[90] Miles Davis, *Get Up with It*, produced by Teo Macero, Columbia C2K 63970, 2000, compact disc, originally released 1974.

or not the reverb was added by the engineer is difficult to tell. There is definitely artificial reverb used with the electric guitar that affects its timbre, but not, interestingly, its spatial distance from the listener (most likely because this particular electronic effect, along with distortion, has been thoroughly incorporated into the vernacular of electric guitar playing); in fact, in spite of the reverb effect, the guitar seems quite close, on the left.[91] The rhythm guitar and bass guitar parts sound clear and distinct on this track, signaling proximity, while the drums are diffuse, as if recorded at a distance, though slightly shifted to the right of the stereo field. The conga, however, also sounds close and does not exhibit the long reverberation times characteristic of the other instruments. When the flute enters at the 2:00 mark, it sounds highly processed: it has an exaggerated degree of reverberation relative to the direct sound it produces, a balance that does not index any naturally occurring (i.e., non-electronic) recording environment. Yet in the context of an electronic jazz, pop, or rock ensemble, an acculturated listener can easily understand the flutist to be copresent with the other players—we still have the necessary components to understand the recording as recreating a live performance event.

The end of the song form contains a two-bar turnaround phrase that leads back to the top (examples are at 3:27 and 3:59). During these two measures (excepting the downbeat and the pickup back to the first bar) all the players are tacet. The sound that emerges during this tacet moment is what a producer might call room tone: the sound of the space in which the recording takes place, but without any actual playing going on. It is not silence—clearly there is sound present—but rather a neutral background, like a primer

[91] "Electric Guitar | Grove Music," accessed May 24, 2018, http://www.oxfordmusiconline.com/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-1002256412.

on a canvas. Reverberant sound hangs in the air mixed with the buzz of amplifiers, a reminder that it is the musicians themselves who have gone silent and not the recording machinery. An air of expectation is created, partly as a result of the form: turnarounds are designed rhythmically and harmonically to lead back to the top of the form via dominant harmony, and a full stop by all the players can intensify the phrase's dominant function. But the room tone that emerges in the break indicates an actual performance/recording event and the space in which it occurred. Significantly, the reverberation indicates a *place*—a specific locale wherein actual musicians had been playing and are about to play again.

This appearance of a sonic fingerprint of the recording studio connects this track back to the early gramophone recording "Morning on the Farm:" the listening environment is overlaid with the sound characteristics of a foreign place. In terms of Hansen's mixed reality, we have two "channels" of space to access: our own listening environment and the recording environment conveyed by the album track. When there is a sudden change in the reproduced space and the object of our attention is overtaken by room noise, the sense of irreality is heightened, and the contradiction of two spaces coexisting is brought into relief. The room noise of the studio momentarily covers over the room noise of our own listening space, bringing about a greater degree of Hansen's seamlessness between the two sound worlds. We aren't transported to the performance because that room tone now comprises our room tone. It colors our perception of our own space in a way that being "transported" does not. Rather than effect escape, it solicits the use of the listener's productive imagination in order to apprehend the moment. Again, this is not the verisimilitude of VR that transports us elsewhere, but the deployment of a mixed reality that functions to bring

these two sound worlds into contact, and with an integration that facilitates the reimagining of our own space and our position within it.

1.9 Creating Space by Subverting Place: Miles Davis, "Rated X"

Playing music recordings can be likened to the virtual reality artworks that Hansen considers. His term "interpenetrate" works well to describe the mingling of sonic realms that arise while listening to music recorded in a studio that can take advantage of a host of recording and processing techniques to deliver spatial experiences not easily achieved in live, acoustic (i.e., unamplified) performance venues. As noted in the discussion of "Maiysha," it is not the realism that is uncanny, but the awareness of our own space that comes about when reconciling competing arenas for exploration. Because Hansen's theory claims these different arenas interpenetrate one another within our own environment, it is a model that can help account for the phenomenological experience of listening to recordings such as Davis's that highlight the improbability of the spaces which they are purported to represent.

We have just seen how playing recorded music can engage a listener in such a way that separate "channels" of information can blend together through the body's ability to accommodate discrete sonic landscapes. This enables the music to affect the listener's space through an activity of listening that is concerned with making sense of what they hear. The reorientation that is necessitated in those moments when the music manages to throw us off balance (or, to use Merleau-Ponty's visual description, when our world is thrown out of focus) is the source of our renewed perspective on our situation. The tacet

moment of "Maiysha" makes it difficult to shunt room noise, and hence the room itself, into the background.

Changes in the recorded environment, disruptions in the depiction of a fixed performance space, have dramatic effects on how we perceive not only the space around us, but also the quality of the music within it. How much space music takes up, how it fills the space, its *volume*—not in terms of amplitude, but of fullness—can be affected by processing in the studio during recording and post-production.[92]

"Rated X," track 4 on *Get Up with It*, manages to construct experiences of absence and presence at select moments interspersed throughout a very densely textured groove-oriented track.[93] At several points in the recording the sound of the rhythm section is switched out entirely, leaving only the harsh whine of the electronic organ, played by Davis himself (for example, 1:40–1:50). Instruments that drop out include guitar, bass, drums and percussion. In his book on Davis's electronic music, jazz critic and journalist Phil Freeman recounts: "The band laid down rhythm tracks, Miles' organ was dubbed in later by Teo Macero."[94] Macero then used a switch on the mixing board to cut out the signal of the combined rhythm track.

During those beats when a broad swath of the music is cut out, the electronic organ continues to sound, playing dissonant clusters and simultaneities that remain unaffected after the other instruments are popped back in. As the switching occurs on the beat, the

92 In the following chapter, we will have the opportunity for further inquiry into the concept of volume in terms of *depth*.

93 The album is available on Spotify: https://open.spotify.com/track/6LXh8NJUgPAnbHCfEj6bwB. "Rated X" can be also be found on YouTube: https://www.youtube.com/watch?v=mrjFtbGKqFk.

94 Phil Freeman, *Running the Voodoo Down: The Electric Music of Miles Davis* (Berkeley, CA: Publishers Group West, 2005), 111.

muted portions function as caesuras of a sort—though brought about quite differently than those in "Maiysha." These muted beats produce brief moments when it seems as if both the temporal and spatial characteristics of the piece are suspended due to the disorientation generated by the dramatic textural contrast. For a track constructed as a representation of a live performance, this switching breaks the frame of traditional jazz recording practice. (In contrast, Stockhausen's *Gesang der Jünglinge* has no pretension of representing a live performance. As noted previously, the opposition between live and recorded music no longer holds: playing the recording *is* performing the work.) And though a different spatial sense is conveyed in the intervening moments by the unaccompanied electronic organ, the break is too short to allow for any reorientation to the new texture. On their own, the organ clusters seem to sound much harsher, louder, and more widely diffused after the bass and rhythm section drop out. But the channels carrying the rhythm section's audio are quickly switched back in and the organ is once again subsumed into the dense instrumental mix.

The silencing of multiple instruments at once is clearly not a result of actions taken by the musicians during performance. Such instant stoppage of sound is not possible in any actual performance space with its attendant reverberant properties. There is, for example, no lingering ring of cymbals or bass that would normally accompany such a caesura. It takes only a modicum of technological sophistication for a listener to recognize this to be the result of studio manipulation of some sort, as it is too drastic to have been accomplished without any electronic intervention.

"Rated X" therefore calls attention to the arbitrariness of the idea that a recording documents a performance—regardless of whether that performance occurred in front of an audience in a hall, with the microphone acting as an eavesdropper, or in a recording studio,

with no audience other than technicians and their equipment. In this sense, the sounds of the performance function similarly to diegetic sound in film: that is, they are sounds that arise as part of a scene. The abrupt muting of the rhythm section, however, is non-diegetic: it is not part of the scene represented by the recording. As such, it expands our frame of reference beyond the portrayed performance event of organist and rhythm section, and we realize there are other forces at work shaping the track's represented performance. We can no longer imagine ourselves transported to an original performance situation, because there no longer seems to have actually been one in any traditional sense.

This break in the frame causes the attention to momentarily hover, trying to "place" the organ sound in relation to the other instruments which have suddenly ceased to sound. We are aware of another level of performance—the engineer in the studio "playing" a mixing board. Yet Davis's organ continues on, a performance now lifted out of the ordinary jazz-ensemble-in-the-studio context. No longer having this studio "channel" to tune into, our frame of reference reverts back into our domestic space, with a playback system bringing us sounds from the outside world. This helps explain the sudden loudness of the organ and its sudden full, enveloping character. It is an outsize sound compared to the space we are now made aware of occupying. These differing perspectives conflict with one another and force us to become aware of our own listening situation. Whether we are engaged in the activity of playing a CD or streaming an MP3, we are in effect occasioning a performance of music ourselves in the moment.

Simon Frith acknowledges this view of listening to records (or listening to any media format) as musical performance. In discussing the highly stylized conventions of rock music recordings, he notes that even though it may be evident that a recording does

not document any actual real-time event, we are still conditioned to attend to it as a performance. However, in documenting his own listening practices, he places the performance event not at some other, imagined place, but in the act of record listening, in the same place and time as the listener. That is, he implies that playing a rock recording *is* a musical performance:

> Just as a singer is both performing the song and performing the performance of the song, so we, as an audience, are listening both to the song and to its performance.... I listen to records in the full knowledge that what I hear is something that never existed, that never could exist, as a "performance," something happening in a single time and space; nevertheless, it *is now* happening, in a single time and space: it is thus a performance and I hear it as one, imagine the performers performing even when this just means a deejay mixing a track, an engineer pulling knobs.[95]

As with Stockhausen's *Gesang*, the "now" of listener and music in a single space affords a hearing that redefines the notion of a musical performance.

Yet we can also imagine the producer playing the mixing board as one might play an electronic instrument, "pulling knobs" that cause sounds to be produced and then reversing the action—just as one might depress, hold, then release a key on a keyboard. But as this "instrument" is being recorded directly to tape, there is no intervening device such as a microphone to detect room noise and reflected sound. No "place" is conveyed along with the sound, though at some actual location during production an unseen operator must be at work. These gaps call attention to the recording process necessary to capture the music, and reveal that process as a non-transparent mediation that can be incorporated into the structure of the work. It is this combination of the discordant frames of staged performance

[95] Frith, *Performing Rites*, 210.

and electronically engineered sound reproduction that is responsible for the jarring effect of "Rated X."

Playing CDs through loudspeakers creates spaces that are shared by both the listener and the music, and perhaps even by virtual musicians who are understood to be producing the music. This idea of a shared space is challenged by disruptions such as those that occur in "Rated X:" they alter the character of the listening space dramatically. The dense, highly active texture that had helped define the space is abruptly cut out. As it is only the organ sounds that remain, its timbre and simultaneities are suddenly highlighted. And by being highlighted, they are likewise altered: they sound louder, brasher, and even more extensive, as if their reach, with other sounds out of the way, were now allowed to extend further than they had before. This is not the actual case in terms of sound energy, but in terms of listener focus and attention, change in texture certainly amounts to an expansion of the organ sound.[96] The spatial alteration that is effected here is not one where the dimensions of the listener's space seem to have altered. It is more as if the organ sound has "spilled out," or has been allowed to overrun its borders, spreading out and enveloping us in a manner that the thick texture of rhythm section inhibits. This augmentation of the sound allows us a glimpse into how space can be generated affectively through music. Recordings do not simply provide us a simulacrum of a different space, but through the act of listening to them, they afford us an opportunity to experience our own space differently.

When a listener plays a CD, they are choosing to engineer a particular spatiotemporal experience for themselves. The music they select, the volume they play it at,

96 Compare Merleau-Ponty's remarks on the appearance of a full moon on the horizon. Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Donald A. Landes (New York: Routledge, 2012), 36.

the stance they take toward it, and the activity and attention they bring to it, all combine to create a spatial experience. For example, when I play a CD, I adjust the volume to fill the room in a certain way, and I adjust equalization so that the bass does not overwhelm the higher frequencies—I want the music to have a particular presence that will provide me the proper distance to both interact with it and to apprehend it. Simon Frith alludes to just such a presence when he describes his experience of a recording as being a performance that is happening "now." At quieter volumes, I may be less likely to engage with recording as a performance—it will still be happening now, but it would be the now of ambient sound and not of a performance. For example, music that might otherwise induce entrainment by the listener might be too soft for that mode of engagement to be solicited—and thus not be attended to as a music performance. At louder volumes, if distortion and objects rattling on my shelves distract my attention, the recording once again becomes ambient sound, and no longer a musical performance. The configuration of the home listening situation is engineered by listeners to afford certain engagements with the recording.

In contrast to "Maiysha," the strangeness and power of the "Rated X" caesuras stem in part from the way in which the organ is allowed to carry on untouched by the sonic blackout of the other players. The status of the organ's location within the recording's sound stage is no longer clear, as the sound stage itself has evaporated. Macero has made the organ placeless: disrupting the "live performance" convention of jazz recording. The organ is actually more like the sounds produced by Stockhausen's *Gesang* or by a hi-fi aficionado playing a sound effects record: a non-domestic sound brought into the home (and amplified). When the rhythm section disappears, the competing frames of reference

cause a moment of disorientation that further denaturalize the spatiality of the listening situation, prompting a reassessment of the surroundings and the sounds within them.

Even when it seems as if the listener is a passive participant, subjected to the whims of the audio engineer, the listener still has an active role to play in reframing the event. Otherwise, the music-listening activity is in danger of ceasing altogether. Listening to recorded media is an active process that needs to negotiate these competing channels of sound and the contexts they bring with them. And that means acknowledging the listening situation itself, being situated within a listening situation. These disruptions in "Rated X" are solicitations by the music to tune into this effect. And as they involve manipulations of the medium itself, they require adjustments that bring the actual listening space back into awareness. Such adjustments can be shown to involve the body if, for example, the listener were entrained with the pulse of the rhythm section, tapping their foot or nodding their head. When there is no longer a pulse to entrain to, the tapping stops and the body is no longer engaged with the music in the same way and must seek another approach. (Of course, the tapping could continue, and would also be a mode of engagement with the music, but it would not be entrainment.)

Somehow, an equilibrium must be established for the activity of music listening to continue, however uneasy and provisional. In Merleau-Ponty's terms, equilibrium is the balance sought by a perceiver's body in its attempts to achieve an "optimal grip" on an object—neither too close nor too far—that will allow it to best be viewed.[97] This grip-seeking behavior holds for listening as well. A reevaluation of the listening situation in order to gain a purchase on the music is precisely the type of productive work "Rated X"

97 Taylor Carman, *Merleau-Ponty* (New York: Routledge, 2008), 110.

affords. It serves as an example of how music is able to offer opportunities and alternatives for listeners to remake their own worlds through the presentation and creative manipulation of indicators of space and place.

The adjustments to the electronically mediated changes in texture in both of these Miles Davis tracks point out that the listener forges a spatial connection to the music during the playback of a recording. In "Rated X" the rhythm section's disappearance—and the resultant feeling of its absence—is the very thing that points out the fact that the section had some substantive existence in the first place. This is evidence that there is no neutral position from which a listener might perform an evaluation of the music. Listening is always a contextualized, intentional activity and music is always accessed affectively. This is because perception entails engagement by an embodied listener, motivated by needs and desires, both of which are affective concerns. Consequently, if perception is never disinterested, we cannot expect aesthetic appreciation to be either.

So much of what we hear when listening to recorded music is an artifact of the recording process itself, yet there are great troves of music that we know *only* through recordings. Therefore, there is a lot we can miss when we don't take recording *and* playback into account. Concert listening, home listening through loudspeakers, personal listening through headphones and earbuds, each of these can focus the listener's attention on different aspects of musical structure. And each can give different affective experiences of the music. It is in the listening event, therefore, that some of music's most distinctive characteristics are revealed. And, as the history recounted above has demonstrated, that event is predicated upon a set of conventions for music listening and the possibilities (and constraints) of recording technology. In the following part of the dissertation I will draw

upon these contextual framing devices and our affective responses to the sounds we hear to describe other spatial experiences afforded by music—and by recorded music in particular.

PART II

Chapter 2: Topographical vs. Phenomenological Space

2.1 Definitions: contrasting concepts of space

The first part of this dissertation discussed how space in music is often understood in topographical terms, as points of sound plotted out in relation to a three-dimensional coordinate system of constant units. Space so considered is an empty container filled with localizable objects at measurable distances from one another. Musically, this means that sounds have points of origin that can be located relative to each another and to a listener in the same manner as objects laid out schematically in a blueprint. This conceptual framework takes advantage of a tendency to objectify the things we encounter in our surroundings; that is, to consider things we observe as having an existence independent of our own perceptual processes.[98]

But there is more to space than the side-by-side arrangement of things. Along with this juxtaposition of objects, we also experience depth. Depth is not the measurable distance between objects. It is the distance between ourselves and those objects. The space as container model cannot account for ourselves as listeners because although we may be surrounded by objects, we are not enveloped in the same space as them. According to Merleau-Ponty, "depth is the dimension according to which things or the elements of things

98 This tendency is attested to by Merleau-Ponty. He considers this ability of consciousness to forget itself to be the origin of what he calls "objective thought," a kind of pretense of consciousness. In objective though, we forget that perception is perspectival and are drawn into a belief in the autonomy of the object: "our perception ends in objects and the object, once constituted, appears as the reason for all the experiences of it that we have had or that we could have." This occurs due to the stability and unity provided by the horizon (i.e., the background) that surrounds it—ensuring that the object exists beyond just the side it presents for viewing. Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Donald A. Landes (New York: Routledge, 2012), 69–70.

envelop each other, while breadth and height are the dimensions according to which they are juxtaposed."[99] Or as Kozak explains, the problem with a model of space based upon juxtaposition is that "depth, rather than securing its own status as a unique presence in our experience, becomes equivalent to width viewed from a different angle."[100]

In this container scenario, all unique features of the landscape are flattened out, interchangeable: Kozak continues, this is the view of "an outsider who has no access to the objects that inhabit the space." As a result, there is nothing more we can learn about the object. Once we know its position, once we have posited the sides not seen, once we attribute the fact that one object standing in front of another occludes our view is not due to our own perspective but to the locations of the objects, depth disappears. Objective space is comprised of the distances of objects from one another. But depth is different. In depth, things are revealed or hidden depending upon where we are situated.

This container concept of space is a symptom of what Merleau-Ponty calls "objective thought," a mode of thinking about our surroundings that is secondary to perception. Consider the Miles Davis examples presented in the previous section. Objectively speaking, Davis's organ part does not change in amplitude, nor is there any measurement of the volume of space it fills that can distinguish the portions of the track without the rhythm section from those portions with it. In order to understand the musical effect of the organ's increased loudness and presence, we have to appeal to affect.

By contrast, we can consider a phenomenological, inhabited space, one which is similarly structured through relative positions, but does not correspond to any external

99 Merleau-Ponty, 276.

100 Mariusz Kozak, "Listeners' Bodies in Music Analysis: Gestures, Motor Intentionality, and Models," *Music Theory Online* 21, no. 3 (2015), http://www.mtosmt.org/issues/mto.15.21.3/mto.15.21.3.kozak.html.

standard of measurement. The container concept overlooks the experience of space that is not simply occupied by sounding objects, but by sound itself. For example, a standing wave can only be heard in certain parts of a room and not in others. In this case, the sound itself, not just the sound source, occupies a location. A sound can affect our bodies through soliciting a movement or posture to take it in, thereby altering our relation to the space around us. Empty rooms, full rooms, dark rooms, light rooms, quiet rooms, loud rooms—all of these provide experiences of space that are not simply the result of objective quantities or emotional reactions. Space needs to be inhabited by a perceiver for such room qualities to be perceived. Simply taking measurements won't do. Similarly, in order to understand the musical effect of Davis's suddenly unaccompanied organ clusters, we have no choice but to appeal to affect.

David Morris considers the concept of envelopment to be central to how we perceive depth. He specifies a logic of envelopment based on the idea that "parts of things" envelop one another in accordance with the way body parts "envelop one another in movement," and, importantly, that this envelopment necessarily occurs within, and is constrained by, a larger space. "Our movement with the world is folded through this topology and thence expresses a *sens* of depth."[101] We experience qualities of depth and volume in space that are not entirely congruent with objective measurements of size and location. This is because ours is a lived space that must continually be negotiated by the body. Hence, it is labile; our efforts in search of the "maximal view" or optimum perspective attest to this shifting quality.

101 Morris, 108. Morris here uses the French term *sens* as he deliberately draws on the different meanings of the word: as "sensory-perceptual experience," as "meaning," and—in the specifically French usage that warrants the untranslated, italicized expression—as "direction." Morris, 24.

Morris provides a pair of examples that illustrate that what we sense as space comes down to a matter of our situation, rather than our position.

> In the film *The Thin Red Line*, there is a very long sequence in which the camera dwells on a sea of grass cloaking a hill the soldiers are to occupy; the camera moves as if an eye transfixed by every swaying movement of the grass. This moving absorption conveys a sense of an immense distance up the hill; the distance would appear much smaller if the camera approached the hill as someone looking at its top or over it, as someone not possessed by what might be hidden in the cloak of grass. My being late and far from my destination provokes a similar but converse effect: in my haste I do not linger over details spread within the plaza, the spread of the plaza slips into the background of my perception, and correlatively the plaza seems to contract. Again, a city block usually appears longer if one is dawdling and window-shopping, absorbed in detail, as opposed to driving through the block, treating it as a thoroughfare.[102]

These examples show how depth appears differently in different situations, situations that involve the affective state of the perceiver.

Part I described musical works that made use of an objectified concept of space. Although the studio effects of "Rated X" are not ones that can be produced by musicians themselves during a concert performance, they index just such a context through disrupting it. Similarly, Stockhausen's use of multichannel recording technology in *Gesang der Jünglinge* clearly calls attention to the location of the sounds in his work. By contrast, the spatial aspects of the works considered in this part of the study do not always have such easy markers of spatiality. The analyses are offered as possible hearings—hearings that are not only plausible, but, ideally, likely ones—bringing to light observations and feelings that the listener may not have been aware of, though had experienced nonetheless.

102 Morris, *The Sense of Space*, 17.

Because I am describing an experience of a listener, rather than embarking on a study of performance practice or doing a score-based analysis, I will follow the principles of phenomenological analysis set forth by Marion A. Guck. She argues that in order to understand music, "we must investigate how music and listeners interact."[103] Though this interaction is personal, Guck claims that it can nevertheless be intersubjective. This is achieved by "crossing back and forth between layers of intentional objects and states" in order to "identify the content of the objects or states while also ensuring the integrity of the musical object."[104] I will also compare my observations with those of other phenomenologists of music, and demonstrate how perception of space is a concern for other theorists as well, even if not always addressed as such. I can therefore maintain that the hearings presented here are not entirely idiosyncratic; the work on perception, affect, and meaning that informs these analyses points toward a spatial component that is part of how we make sense of the world around us.

Each of the analyses in this section focuses on different senses of space we can perceive in musical structures such as harmony, texture, and timbre. They also demonstrate how certain musical structures that might be dismissed as extramusical play a key role in determining the ultimate spatial effect of any piece of music: for example, the environment in which a recording was made and the reverberant properties it contributes to the music, or the environment in which a recording is played back and the ability of walls and floors to reflect sound waves and transmit vibrations. I will also consider how

103 Marion A. Guck, "Analysis as Interpretation: Interaction, Intentionality, Invention," *Music Theory Spectrum* 28, no. 2 (2006): 194.
104 Guck, 202.

timbre is inherently imbued with characteristics that provide all music with the ability to shape our spatial experience.

The first analysis is of a moment in the Sanctus from Johannes Ockeghem's *Missa mi-mi*, where a four-voice polyphonic texture changes to the resonant sonority of a sustained triad. This change illustrates one of the most basic elements of Morris's concept of voluminosity: enveloping actions take place within a larger enclosing space. Therefore, if a musical object is suddenly heard as more voluminous, the spatial situation will necessarily become enlarged to accommodate it. Having established the basic principle that music affects our perception of space, I consider a single movement from George Crumb's *Black Angels*—II: Absence—and discuss how its timbral qualities direct our awareness to different aspects of our surroundings and thereby contribute to the piece's expressive qualities. I focus in particular on the sense of immersion and how it can arise from acoustic properties of an instrument's timbre. I then turn to an electronic work by Éliane Radigue and demonstrate how an embodied approach to listening is necessary in order to account for its distinctive textural quality. Finally, I draw on the conclusions made through these analyses to provide a hermeneutic account of the first movement of Steve Reich's *Electric Counterpoint*. Grounding a hermeneutics of music on the perspective of an embodied listener shows how interpretation can be based on principles other than the personal narrative associations of an individual listener and thus serve as an intersubjective basis for critique and analysis.

2.2 Merleau-Ponty and Gestalt Psychology

The work of Merleau-Ponty gives us insight into music's affective capacity by insisting on a theory of perception as a felt phenomenon, rather than a theory of sensation as a detection mechanism (akin to an electronic motion detector that can open a door automatically when someone approaches it). For Merleau-Ponty, this latter model is insufficient to explain how some actions are carried out independently of conscious thought, and instead seem to be situated, and even embedded, within a body habituated to particular activities. Merleau-Ponty was fascinated by a 1920 psychological case study by Adhémar Gelb and Kurt Goldstein of a patient they refer to as Schneider, a man with a brain injury that left him unable to name or recognize ordinary objects that he nonetheless was able to handle and use appropriately. This curious disparity between cognitive and physical functioning led Merleau-Ponty to the conclusion that the body is a more primordial organizer of sensation than any purported representational image in the mind. And as the body is situated in an environment, with kinesthetic relations to the objects within its purview, Merleau-Ponty surmised that a concept of space is built into our very existence. Merleau-Ponty writes, "we have to reject the age-old assumptions that put the body in the world and the seer in the body, or, conversely, the world and the body in the seer as in a box."[105] Instead, perception is an active, embodied engagement with the world, or as

105 Maurice Merleau-Ponty, *The Visible and the Invisible; Followed by Working Notes* (Evanston, IL: Northwestern University Press, 1968), 138.

philosopher Scott Marratto writes, "we perceive things as eliciting the movements of our bodies."[106]

Merleau-Ponty's theories are closely aligned to those of Gestalt psychology (and, not incidentally, to subsequent Structuralist theories as well) that posit a "diacritic" structure to cognition. Diacritic is a term Merleau-Ponty uses often in his work. It refers to the organization of phenomena into pairs of opposing, mutually defined categories (an organization that is the very basis of the Structuralist enterprise). The Gestalt pairing of "foreground" with "background" is a prime, and indeed foundational, example.

The discoveries of Gestalt psychology formed the basis of Merleau-Ponty's phenomenology:[107]

> Gestalt theory informs us that a figure on a background is the simplest sense-given available to us.... It is the very definition of the phenomenon of perception, that without which a phenomenon cannot be said to be perception at all. The perceptual 'something' is always in the middle of something else, it always forms part of a 'field'. A really homogeneous area offering *nothing to be* cannot be given to *any perception*.[108]

In his later work, Merleau-Ponty went so far as to claim that "the figure on a ground, the simplest '*Etwas*'—the *Gestalt* contains the key to the problem of the mind."[109]

A background is necessary for a foreground to appear. Background and foreground are mutually exclusive yet complementarily defined: whatever is not in the foreground is in the background, and vice versa. This simple division is evident in other domains, especially

106 Scott L. Marratto, *The Intercorporeal Self: Merleau-Ponty on Subjectivity* (Albany: State University of New York Press, 2012), 32.
107 Lester Embree, "Merleau-Ponty's Examination of Gestalt Psychology," *Research in Phenomenology* 10 (January 1, 1980): 90.
108 Merleau-Ponty, *Phenomenology of Perception*, 4. (Italics in original.) I will have occasion to say more about homogeneity, texture and perception in later chapters.
109 Merleau-Ponty, *The Visible and the Invisible; Followed by Working Notes*, 192.

language, where good and bad, near and far, large and small, form similarly constructed binaries (as do, in certain contexts, dog and cat, or walking and driving—arbitrary cultural antonyms that likewise provide schemas for organizing perceptions). The idea here is not that we view everything in a crudely delineated either-or, black-and-white scenario, but that we tend to identify things by contrast to other things: the defining feature of the Saussurean theory of language. Merleau-Ponty writes at the beginning of "Indirect Language and the Voices of Silence:

> What we have learned from Saussure is that, taken singly, signs do not signify anything, and that each one of them does not so much express a meaning as a mark a divergence of meaning between itself and other signs.... We may conclude that language is made of differences without terms; or more exactly, that the terms of language are engendered only by the differences which appear among them.... The unity he is talking about is a unity of coexistence, like that of the sections of an arch which shoulder one another.[110]

For Merleau-Ponty, human gestures function similarly. They are defined in relation to a system. Certainly to one another, but in addition, in relation to a background, a background provided on one level by the surrounding environment, but at an even more basic level, by the body and its own capabilities. Marratto explains: "Our gestures are expressive and generative of meaning insofar as they articulate and differentiate themselves according to the same kinds of diacritical structures that govern the relations between signs in any conventional language."[111]

110 Maurice Merleau-Ponty, "Indirect Language and the Voices of Silence," in *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, ed. Galen A. Johnson, trans. Michael B. Smith, Northwestern University Studies in Phenomenology and Existential Philosophy (Evanston, IL: Northwestern University Press, 1993), 76–77.

111 Marratto, The Intercorporeal Self, 8.

Merleau-Ponty sought to describe how it is that the emergence of such structures through human gesture gives shape and substance to our world. It is through our own body—rather than cognition alone—that we determine what aspects of our surroundings call out to us, and how our needs and desires cause us to seek out and perceive some features at the expense of others. It is through this embodied engagement with the world around us that depth emerges, a quality of the environment that is more than simply another axis analogous to height and width (the view, according to Merleau-Ponty, that Descartes presents in *Dioptics*).[112] This means that depth is a quality one experiences, rather than an objective property of the world, or as he puts it, "it is not indicated upon the object itself, it clearly belongs to perspective and not to things."[113]

A difficulty in using Merleau-Ponty's theory of embodied perception in music studies is that a significant subset of musical listening practices privileges an immobility of the body. And Merleau-Ponty claims that it is a body in motion that perceives. However, it is important to note two things here. One is that Merleau-Ponty is opposing a moving body to the idea of an ego or consciousness that is detached from the world, autonomous, that simply takes in information in the form of sense data. He often refers to the potentiality of the body's movements, the awareness of a situation in which possibilities are presented as potentials for action. This is why Merleau-Ponty does not oppose Descartes's "I think" to "I do," but rather to "I can." The world is a situation in which a body finds itself according to its current concerns regarding tasks and goals. Second, concert listening is not entirely

112 Although Merleau-Ponty, rather than demoting height and width in relation to depth, instead elevates those dimensions to a similar existential quality related to the movement of the body. (See Romdenh-Romluc (2011), pp. 109–110.)

113 Merleau-Ponty, Phenomenology of Perception, 267.

immobile. Even the gesture of closing one's eyes is an active, bodily engagement with music, not to mention slight swaying, head movements, occasional deep breaths, contented smiles and furrowed brows—all these are sanctioned and even encouraged as ways of engaging with the music without unduly impinging upon the engagement of other listeners. Therefore, we can continue to explore embodied perception in relation to modes of listening that involve a seated, seemingly impassive listener. Though we can understand the social norms that constrain the motions of a seated listener, the body and its capabilities remain the ultimate ground of all our perceptions.

Kozak emphasizes the fact that the gestures Merleau-Ponty claims are necessary for grasping the world around us are not metaphorical. As such, they provide support for Lewin's transformational music theory: "such gestures are not abstract shapings of some metaphysical musical energy—a kind of symbolic action that an analyst might effect in lieu of *actually* performing the music—but real movements of real, embodied listeners. They are the non-figurative, tangible source of *doing* something to a *Klang*."[114]

Closely aligned with Merleau-Ponty's phenomenology is the work of psychologist James J. Gibson and his ecological approach of perception. Gibson was intent on dismantling a model of perception in which the brain was responsible for parsing and identifying sensory input in order to come up with a mental representation of the world around it. Instead, he proposed that vision is accomplished through motion, by means of an entire visual system, not just an eyeball that acts like a camera: "when there are no constraints, we look around, walk up to something interesting and move around it so as to

114 Kozak, "Listeners' Bodies in Music Analysis."

see it from all sides, and go from one vista to another."[115] In his view, perception is an interaction with the environment, and not, for example, a sequence of snapshots transmitted to the brain. Such rote transmission, Gibson claims, is based on models of communication that involve a message delivered through a medium and decoded by a receiver. He opposes this to his model of "natural" perception that does not necessitate a separate decoding step (and thus, he claims, is "direct"). Because objects are perceived with respect to their "affordances," a separate interpreting, evaluating step is not needed.

The concept of "affordances" is central to Gibson's theory of perception. Affordances are features of the environment that are relevant to an organism's moment-to-moment needs, desires, interests and concerns, in other words, "properties taken with reference to the observer."[116] For instance, a chair affords sitting—as does a coffee table, if seating is at a premium. Similarly, a cave may afford shelter, or it may afford hiding.[117] By relating perception to the needs, desires, and interests of the perceiver, Gibson explicitly links perception to affect, thus providing a link to the capacity music has to shape moods and emotions.[118]

What was revolutionary about both Gibson's "environmental approach" and Merleau-Ponty's phenomenology was that neither admitted of any two-stage process of first receiving sensory input and then decoding it into a representation of reality. For

115 James J. Gibson, *The Ecological Approach to Visual Perception* (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1986), 1.

116 Gibson, 143.

117 Compare Merleau-Ponty's remark that "the mere presence of a living being transforms the physical world, makes 'food' appear over here and a 'hiding place' over there." Merleau-Ponty, *Phenomenology of Perception*, 195.

118 The view of perception as a body's engagement with the world around it goes against the grain of centuries of enlightenment and post-enlightenment thought. Even in the 1940s, Wittgenstein (d. 1951) in his late work, maintained the passivity of perception: "To interpret is to think, to do something; seeing is a state." Ludwig Wittgenstein, Philosophical Investigations (New York: Macmillan, 1953), 212e.

example, do we see a mound of pebbles, or do we see gravel? It depends: are we looking for a skipping stone or paving material? Our affective state is a factor in our perceptions as much as the affordances of the things we encounter. There is no parsing of the perceptual field independent of the perceiver.

We can transfer Gibson's concept of affordances to music and speak of the perceptual experiences that music affords. For example, Elizabeth Margulis claims that repetition can direct our attention to different levels of musical organization: "repeated exposures [to a musical passage] triggers an attention shift from more local to more global levels of musical organization."[119] When confronted with the ongoing repetition of certain musical elements, we tend to shunt them to the background and begin to consider more closely those aspects of the music we normally would consider secondary. In Gibson's terms, we might say that the repetition in the passage is an affordance: it affords us the opportunity to tune into, say, the structural or timbral features of a piece of music.

The Gestalt perspective used by Merleau-Ponty been taken up by music theorists as well. Patricia Carpenter sought to define how a musical "piece" comes to be conceived of as a coherent unit (a "work"). She considers diacritical contrasts as emerging through gesture. Writing of *Ionisation* by Edgard Varèse, Carpenter states that however "revolutionary" its musical materials, "I grasp its shape much as I do the first movement of a symphony by Mozart, as a single unified gesture or motion—an introduction, first and second ideas, climax and release, coda."[120] Matthew Butterfield, building on Carpenter's work, defines a musical object as "a cognitive formation that is responsive to the situational constraints of

119 Elizabeth Margulis, On Repeat: How Music Plays the Mind (New York, NY: Oxford University Press, 2014), 9.

120 Patricia Carpenter, "The Musical Object," *Current Musicology*, no. 5 (1967): 57.

performance."[121] He applies Carpenter's ideas to objects that are discernible at shorter time scales: notes and chords and phrases. He finds they are the result of an "effortless" parsing of an acoustic environment into objects to which we "ascribe a provisional fixity… in order to organize our perceptions."[122] Though Butterfield seeks a method to delineate objects in a temporal stream, his acknowledgement of the provisional nature of our attempts at segmentation resonates with Merleau-Ponty's claim regarding the active, seeking nature of perception.

Kozak explicitly links the formation of these objects with the gestures of an active embodied listener:

> By having a motor intentional disposition toward musical sounds, listeners constitute those sounds as entities that require some kind of action that is constrained by each listener's background, motivations, physical capabilities, and so forth. Such constitution effectively turns sounds into musical objects, rather than these objects being given in advance.[123]

This engagement with musical sound, then, takes place in a space that allows for the movement of a listener. And just as in Morris's examples of the situated nature of voluminosity, whatever intentional action we perform affects how we perceive the space around us.

In spite of Merleau-Ponty's suspicion of the objectifying mode of consciousness, the delineation of objects through gesture is, at the very least, an organizing principle for perception, a "provisional fixity" by which to gauge—among other things—an optimal

121 Butterfield (2008, 333).
122 Matthew Butterfield, "The Musical Object Revisited," *Music Analysis* 21, no. 3 (2002): 332.
123 Kozak, "Listeners' Bodies in Music Analysis."

perspective.[124] All of the above-mentioned authors make the point that we carve out these objects through gesture, and then temporarily grant them an objecthood in our field of awareness that allows for further inspection.

From the standpoint of embodied perception, listening to music is, like all acts of listening, related to and dependent upon a moving body, but not in the same manner as, say, dance music is. This immediately raises a question: if the impetus for repetitive, metered, bodily motion is absent, yet the music nonetheless calls forth a response that implicates the body in the listening event, how might we theorize such a response? The following chapters propose how this can be done, by looking beyond entrainment to the bodily gestures a listener performs while seeking a perspective on the music.

2.3 Additional Perspectives on Music and Space

2.3.1 Attention

The notion that our sense of space is mutable and that music has resources to transform it is an idea alluded to by other theorists—typically in discussions of musical motion, but in other ways as well. Marion A. Guck identifies a moment from the Adagio of Mozart's A-Major Piano Concerto, K. 488 (Figures 2 and 3), that conveys a sense of space through a change in what the music affords a listener in terms of potential movement:

> The piano, in its grace and unpredictability, seems created for each listener and we respond to its attention.... As the piano comes to rest, the orchestra enters.... It feels warm, welcomed. We want it and we want it to go on. Why?

[124] Merleau-Ponty admits that such objectifying thought is a "natural continuation" of perception, though it has a propensity to exceed its own basis in experience, resulting in the "death of consciousness." Merleau-Ponty, *Phenomenology of Perception*, 74.

> The immediate reason, I think is the surrounding space in which listeners move more freely.[125]

The attention that is at first focused on detail in the figuration and voicing of the opening theme suddenly has room to roam. This is offered through a relaxation in melodic complexity, a fuller sonority with the addition of winds and strings, and a relaxation in melodic complexity that allows for a transfer of attention toward longer, hypermetrical groupings of two-bar segments.[126] This redirection of attention is discussed at length in Margulis. Citing the work of David Lidov, Margulis explains that when something is repeated several times, our attention is diverted: the repeated material "cancels out its own claim on our attention and thereby refers our focus elsewhere, to another voice or to a changing aspect."[127] We can extend this analysis beyond repetition to habituation: we similarly need to pay less attention to familiar, predictable material. When a melody follows a more conventional path, it does not require the same level of focus: therefore, the range of musical qualities that can be taken in is expanded.

The soft, intricate piano part constricts the listener in a way the orchestral material does not. As Guck notes, its slight irregularity invites the listener to investigate it more closely, to attend to precisely those movements that do not conform to what we might

125 Marion A. Guck, "Music Loving, or the Relationship with the Piece," The Journal of Musicology 15, no. 3 (1997): 350. Also intriguing here is how the piano pays attention to *us*, echoing the manner in which Merleau-Ponty considers things in our visual field to have the ability to solicit our gaze.

126 Guck, 350.

127 Elizabeth Margulis, *On Repeat: How Music Plays the Mind* (New York, NY: Oxford University Press, 2014), 51.

consider a sensed norm at work guiding the line (in this case, "a fairly easily discerned voice-leading frame"):

> We hear very little repetition of figuration or rhythm…almost everything is lovingly detailed, lovingly reworked. And these particularities keep one listening closely, moment by moment.[128]

Through mode of attention, perspective, listening posture—music summons us to comport ourselves in a particular manner in order to best take in what it presents. The affordances of the music, this play between frame and figuration, explain the later spaciousness identified by Guck. When attention is directed toward sonorous qualities, rather than the mechanics of intricate melodic motions, the voluminosity of the music changes. There is a switch between a listener focusing in, enveloping the music in order to better attend to its motions, and then standing back, being enveloped by a more voluminous sound. In my analysis of the Sanctus from Ockeghem's *Missa Mi-mi*, I focus on a similar moment when the

[128] Marion A. Guck, "Music Loving, or the Relationship with the Piece," *The Journal of Musicology* 15, no. 3 (1997): 349.

space that envelops both the listener and the object expands in response to the listening situation.

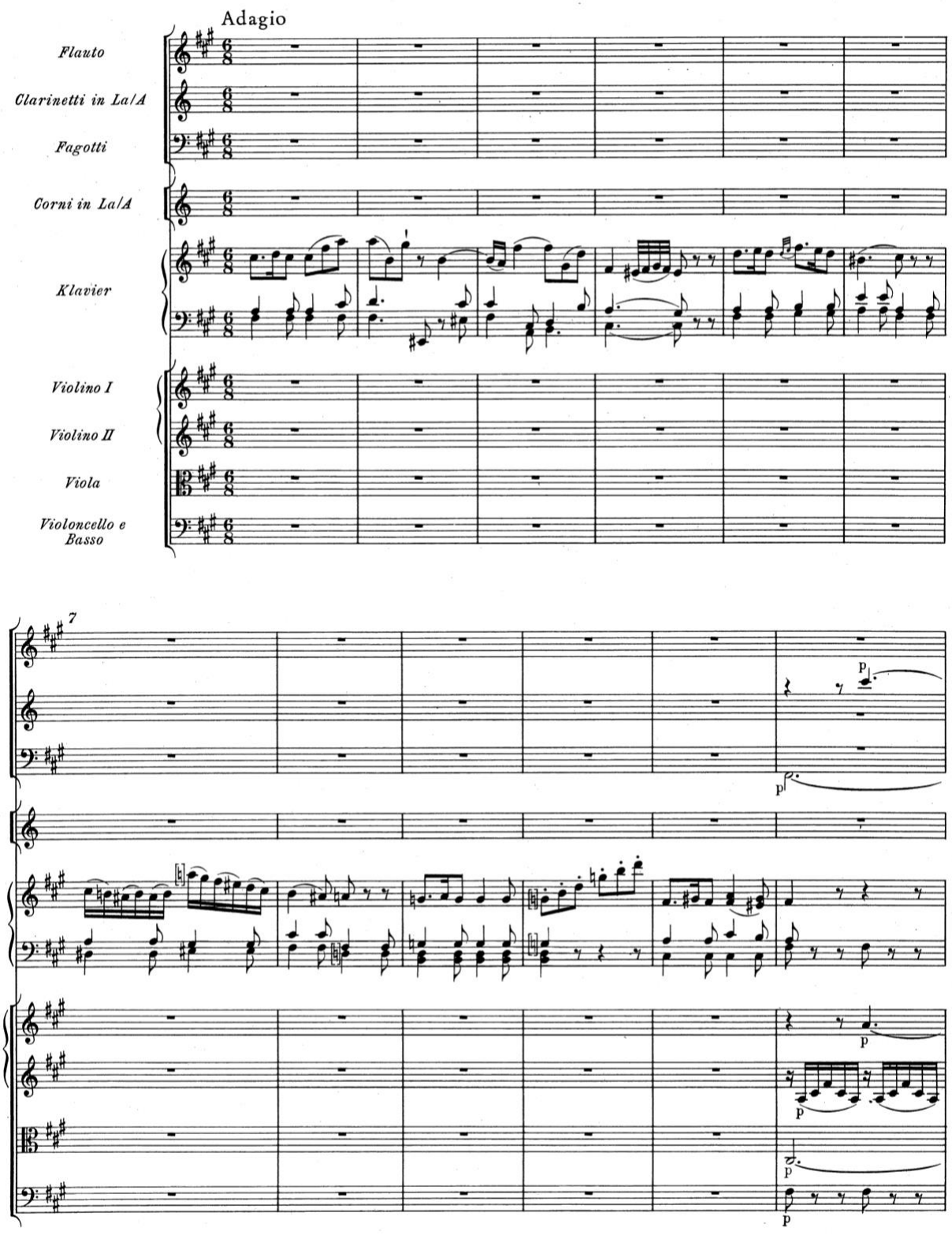

Figure 2. Mozart, Piano Concerto in A Major, K. 488:ii, mm. 1–12.

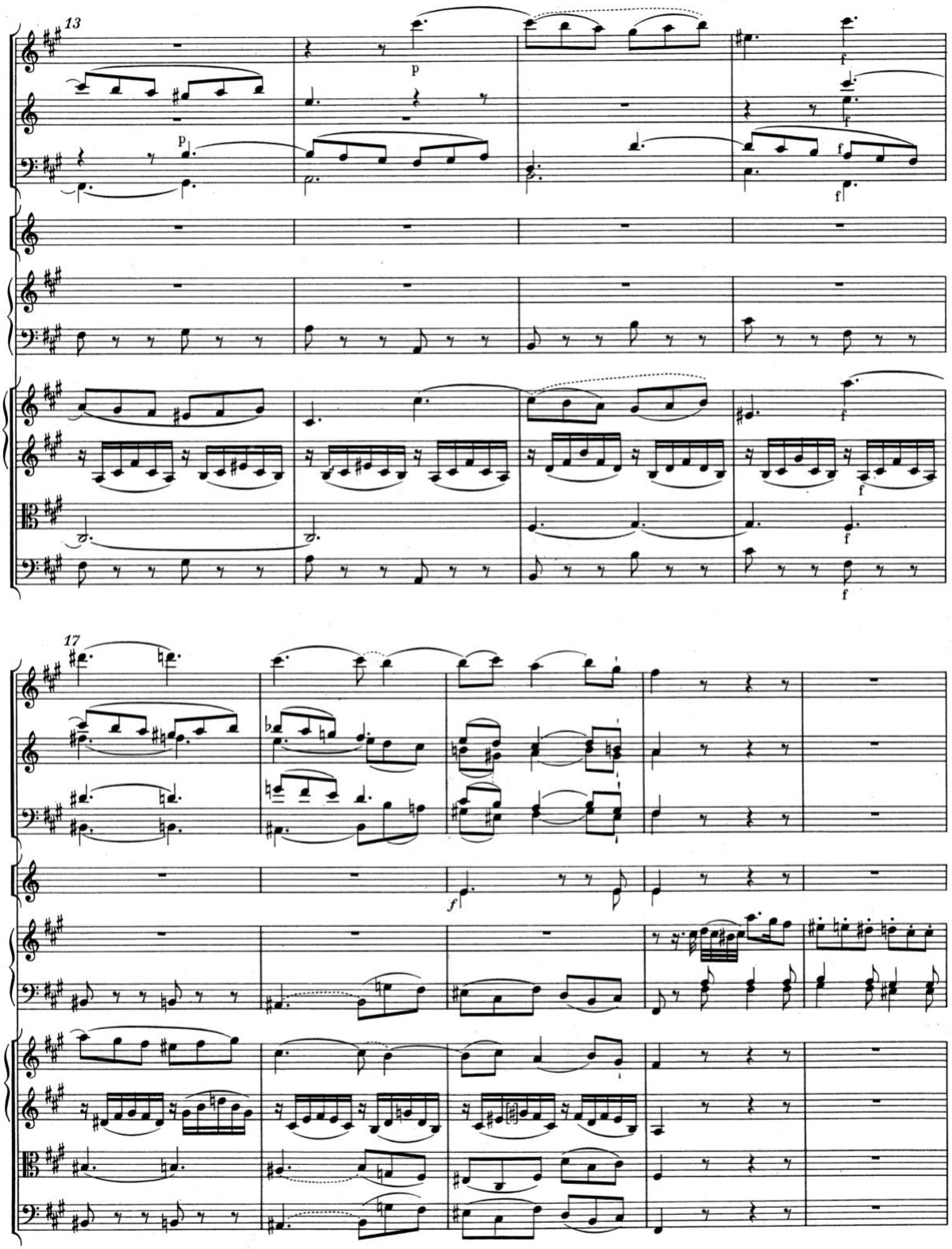

Figure 3. Mozart, Piano Concerto in A Major, K. 488:ii, mm. 13–21.

2.3.2 Musical Propulsion

A similar experience of spatial expansion is detailed by Eric Clarke, though explicitly in terms of motion. The moment Clarke examines comes during an interlude between scenes 2 and 3 in Act III of Alban Berg's opera *Wozzeck* in which a single pitch-class is played by the entire orchestra. Clarke claims that rapid increase in volume in homophonic music specifies self-movement relative to a static environment (in contrast to polyphonic music, which indicates the movement of external objects in relation to one another). Thus, the rapid crescendo throughout the orchestra on the pitch-class B gives listeners the feeling that they are in rapid motion and, moreover, face imminent impact.[129]

Clarke adopts Gibson's approach to perception and uses it to explain the phenomenological sense of movement experienced by music listeners. He concludes that music does not simply convey movement metaphorically or analogically, but that it allows movement to be directly perceived as such. That is, "since sounds in the everyday world specify (among other things) the motional characteristics of their sources, it is inevitable that musical sounds will also specify movements and gestures—both the real movements and gestures involved in actually producing music...and also the fictional movements and gestures of the virtual environment which they conjure up."[130]

129 Clarke, *Ways of Listening*, 76–77. Clarke refers to a section of the track "Build It Up, Tear It Down" (0:28–1:02) from the CD *You've Come a Long Way Baby* (Fat Boy Slim, 1998).

130 Clarke, 74. Fictional," "real," "virtual" are all used here as though readily understood. I have attempted to unpack these terms in Part I and will have more to say about them throughout Part II.

In another example, Clarke discusses an instance of sampling in a dance track in which the sample begins soft and highly processed (through a low-pass filter) and then gradually becomes more clear.

> There is a powerful sense of the source being at first concealed below some kind of acoustic horizon, above which it then progressively rises until it is fully revealed and directly in front of the listener by the time the texture changes at 1:02. The explanation for this strong sense of motion, and for the particular style of motion described here, comes straight from ecological principles: in everyday circumstances, the acoustic array from a sound source that is concealed in relation to a listener (behind a large object like a wall or a building, or below a horizon) will possess attenuated high frequencies due to simple masking principles. High frequencies are absorbed and dissipated in the environment more rapidly than low frequencies, leading to the characteristic "bass heavy" quality of amplified music heard at a distance (the "open air pop festival" sound). As the distance to the source decreases, or the degree of occlusion declines, the high frequencies increase in relative intensity, shifting the timbral balance towards increased brightness.[131]

Here, Clarke explains how a habituated knowledge of the manner in which our own hearing functions (within our own bodies and the stances, attitudes, and positions they take up) allows us to determine more about the sound than a simple frequency analysis would be able to show. The timbres we hear, as measured by the variations in amplitude across the frequency spectrum, is often a function of our location in relation to a sound source. It is therefore understandable that modulations of these amplitudes in music can elicit different spatial sensations in listeners. To continue with the Gibsonian approach, we can say that we are able to sense that our potential for motion has been changed when presented with a different set of affordances in the music.

131 Clarke, 81.

Clarke cites such expressions as "swept away," "transported," "blown away" as evidence of the experiential reality of sensations of self-motion[132] And indeed, such sensations are grounded in actual perception and not to be dismissed as simply rhetorical flourishes used to convey powerful affective states. We draw upon our knowledge of how sounds behave in order to interpret the musical sounds we hear, and this in turn results in not only perceived motion but also a concomitant perception of a space in which that motion can be accommodated. Moreover, this knowledge about the sounds in our environment is more than just a matter of distance or directionality. It is recognition of what the sounds afford with respect to our own efforts to grasp them.

In Clarke's *Wozzeck* example, the listener is said to be propelled through space by the music; in the final section of the Ockeghem Sanctus, it is the listener that is held stationary while the triad spreads out to meet them. We are not pushed through into a new realm, but rather it is the environment in which we are situated that is altered. As Mark Hansen might say, there has been an effortless move from one realm to another within a mixed reality—an effortlessness that speaks not only to the sophistication of the recording technology, but also to the skill of the composer in crafting such an affordance for spatial reorientation. Rather than presenting the variety of space we experience when we are in motion, Ockeghem's four-voice triad generates space by keeping us still.

2.3.3 Embodiment and Standing Still

The German musicologist Heinrich Besseler attempted to describe a more physical,

132 Clarke, 75.

active, non-cerebral engagement with music that was distinct from concert hall listening practices by describing some of music's practical uses. Besseler's 1926 essay "*Grundfragen des musikalischen Hörens*" is an exploration of the characteristics of *Gebrauchsmusik* that separate it from the tradition of European art music, in particular, from the cultural practice that distinguishes a small corps of elite listeners and a performance practice based on the *Werktreue* ideal.[133]

In the decline of the concert hall as the focus of a single musical culture, Besseler saw the beginnings of a relativism that allowed for the co-existence of a multiplicity of styles, all with the value and relevance formerly reserved solely for Viennese classicism and nineteenth-century German romanticism. He explained his perspective on the current state of affairs in high musical culture by tracing a historical succession of Western styles. What is of interest here regarding the Sanctus of *Missa Mi-mi* is what Besseler had to say about homophony.

In the course of his essay, Besseler provides an account of how different styles of music solicit different types of listening. As his concern is with a society's musicking practices, he focuses on the relationship between music and the individual listeners, audiences, and performers who hear it. Like the other theorists discussed in this chapter, he speaks of a type of listening that figures music as an object and of a distance between this object and the listener.[134] He has clearly put his finger on a phenomenological difference between different modes of listening. In addition, he identifies these modes of

133 Heinrich Besseler, "Appendix: Fundamental Issues of Musical Listening (1925)," trans. Matthew Pritchard and Irene Auerbach, *Twentieth Century Music* 8, no. 1 (March 2011): 49–70.
134 Besseler, 64.

listening with different musical styles, which he claims involve different spatial relationships between the listener and the music.

Besseler states that the harmonic objects formed through homophonic compositional practice create a distancing effect, especially when compared to structures created through polyphonic techniques. "Regular fifth and octave sonorities, plus a solostic top voice," he writes, "point unambiguously to an outside listener."[135] This configuration of voices, he claims, occasions a shift in the listener's relationship to the music:

> One comprehends the music as something like an object. The previously direct connection [in the early French motet] between it and the listener [i.e., the listener/performer] has been dropped, and a distance has been instituted.... One has to 'keep still' in order to perceive the music's new and irregular structure.[136]

Besseler, like Guck, recognizes that the type of music we hear is a factor in the stance we take toward it. One type of stance can be keeping still. But this keeping still is not a passive state—it is an active process as the listener seeks an optimum grip on the music. In this case, that grip is best achieved by remaining still.

135 Besseler, 64.
136 Besseler, 84.

Chapter 3: Johannes Ockeghem, *Missa Mi-mi*, "Sanctus"

3.1 Introduction

In this chapter, I will focus on a moment where a voluminous musical object emerges that alters the listener's perception of the size of the surrounding environment. The passage occurs in the Sanctus of Ockeghem's *Missa Mi-mi*, a fifteenth-century polyphonic choral work for four voices, when a two-part contrapuntal texture changes to a four-part sustained harmony: a root position major triad.

In order to describe the spatial effects of this moment I will again draw upon the work of Merleau-Ponty. His description of perception as a type of active behavior (as opposed to a passive reception of sensations relayed to and then decoded by the brain) inserts the body and its kinesthetic abilities into any aesthetic encounter. Aesthetic perception then becomes, as Mark Hansen notes, a matter of framing, initiated in order to take in the object of interest. Combining the work of these two philosophers results in a concept of perspective that is accomplished by our situated bodies through the process of framing.[137] And as this framing process is motivated—that is, pursued in accordance with the needs and the desires of the perceiver—it is, as noted in Part I, affective. Framing allows us to attend to a select portion of the world we want to focus on. Such framing relies on insights gleaned from Gestalt psychology and Merleau-Ponty's interpretation of them, as discussed in Chapter 2.

137 The distinction here between situation and position is Merleau-Ponty's. What makes perspective able to pick certain objects as standing out among others is the perceiver's situation, which is much more than simply a plot point in three dimensions.

3.2 Analysis

The Sanctus from Ockeghem's *Missa Mi-mi* contains sonorities that illustrate how we can hear music as inhabiting a space with a voluminosity that is not simply a function of amplitude (Figures 4–6).[138] At the end of its middle section (measure 47), the piece switches from a vocal duo to a four-voice homophonic texture, and—with the introduction of a triad on the word "Osanna"—immediately changes the affordances of the music. Before the triad, the paired lines, moving in tandem and then in imitation, call upon us to carve out a slice of our surroundings and home in on this twisting motion between the voices. Such a structure solicits a type of framing of space, a determination of what is foreground and what is background. With the triad on "Osanna," our attention is redirected to the resonant qualities of harmony—the voices fusing into a single sound—with consequences for how we grasp the relationship between the music, our environment, and ourselves.

To analyze music phenomenologically requires not only consideration of a performance but of a listening event as well. The following description of the Sanctus is drawn from an experience of listening to a 2007 recording of this work by the Clerks' Group, at home, through loudspeakers, at a moderate volume.[139] Affective features that are accessible in this performance-listening configuration may not be available when listening through desktop speakers or headphones, or even in a concert setting. Yet, the point I wish to make here is that there *are* spatial perceptions afforded by the music, without

138 Measure numbers are taken from Heinrich Besseler's 1930 edition: Johannes Ockeghem, *Missa Mi-Mi*, ed. Heinrich Besseler (Wolfenbüttel: Möseler Verlag, n.d.). Besseler's edition is based on three manuscripts held in the Vatican Library—Chigi C VIII 234, Capp. Sist. 41, and Capp. Sist. 43—and is transposed up a minor third from the source texts.

139 The Clerks Group, *The Ockeghem Collection*, Gaudeamus CD 550/5. Available on Spotify: https://open.spotify.com/track/2nVH5IsN7eQG9TDjVtD854?si=7yudHW9nRySP46bXbd5KNA.

necessarily committing to any one perception that should appear in every listening event. As noted above, what ultimately appears to a listener is a function of that listener's situation. The situations I describe are concerned with musical sound that can change in response to listener movement. This is notably different from listening to music through headphones. In headphone listening, what we hear is not altered by our body movements. The type of mixed reality this affords is a dramatically different one: our aural perspective is fixed. We still interact with such music in an embodied fashion, but often with an awareness of how our perspective on the music does not match up with our own proprioceptive sense—or perhaps, with an awareness of how we *ought* to be moving in order to make the disparate perspectives mesh. Even with eyes closed, headphone listening is an intensely multi-modal phenomenon, as we grapple with the knowledge that—beyond the acousmatic quality of the music, with its hidden sound sources—our body, however it moves, cannot impact that sound. Such music, then, can't truly exist in the same space we inhabit, which likely explains why we typically perceive such sound as being within our own head.[140] When we hear sounds that seem to indicate large, resonant spaces, it is the space in our head that seems to become larger, not our environment. Therefore, as in Part I, the Sanctus analysis, along with the other analyses in Part II, will be concerned with listening to a recording through a pair of loudspeakers.

140 A phenomenon known as "in-head localization" or "intercranial lateralization." Bosun Xie, Head-Related Transfer Function and Virtual Auditory Display (Plantation, Florida: J. Ross Publishing, 2013), 279.

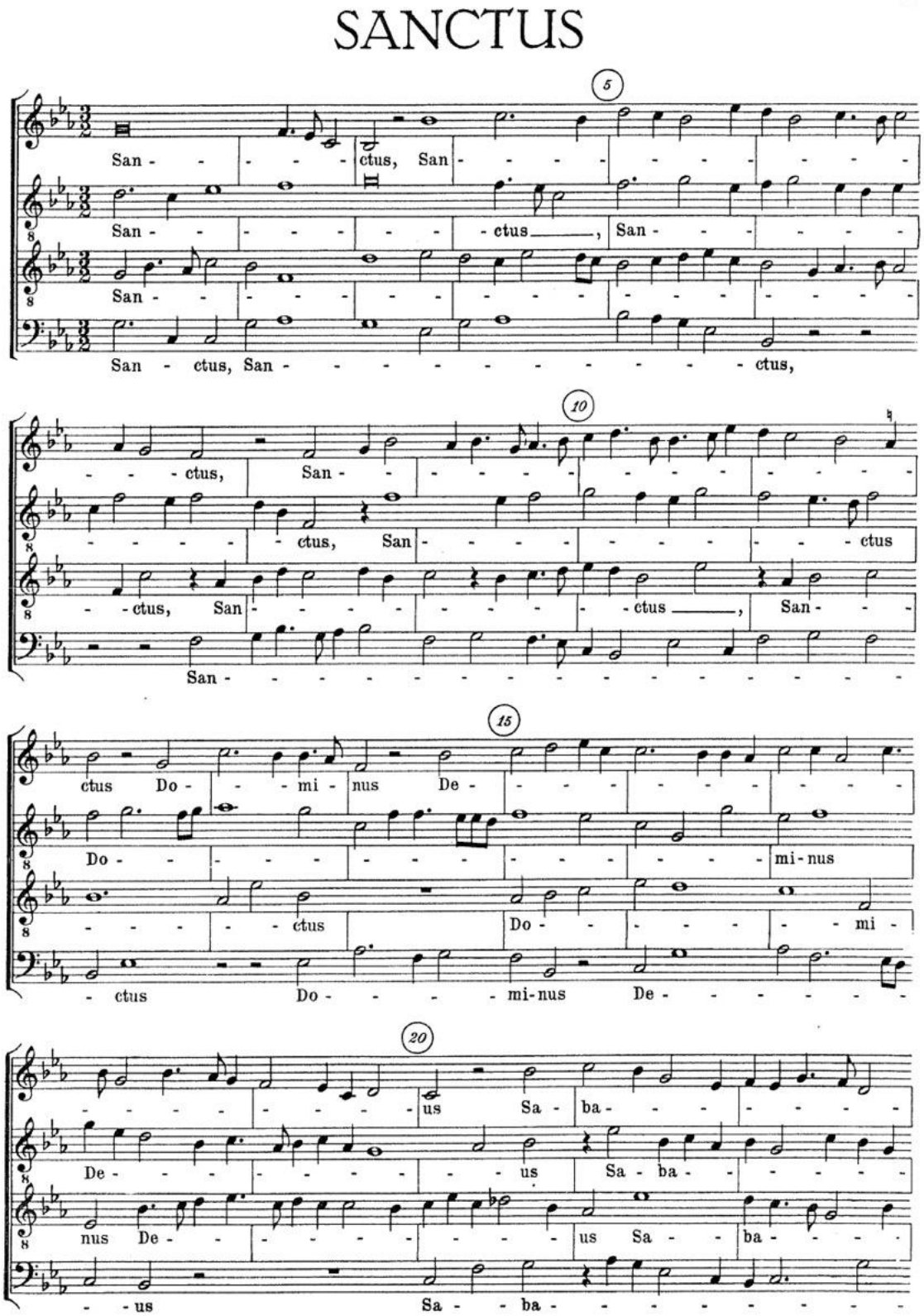

Figure 4. Johannes Ockeghem, Missa Mi-mi, "Sanctus," mm. 1–22

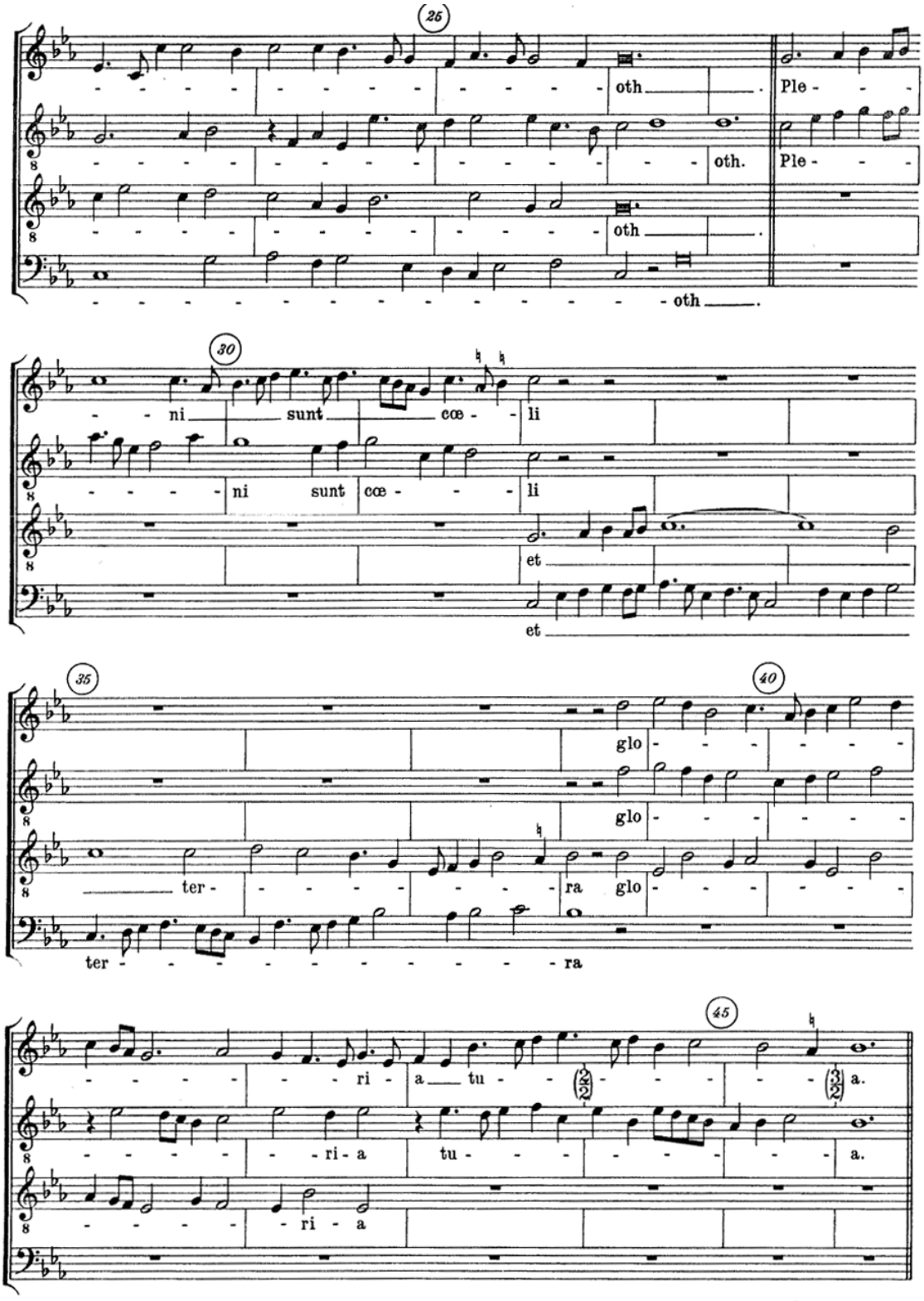

Figure 5. Johannes Ockeghem, Missa Mi-mi, "Sanctus," mm. 23–46.

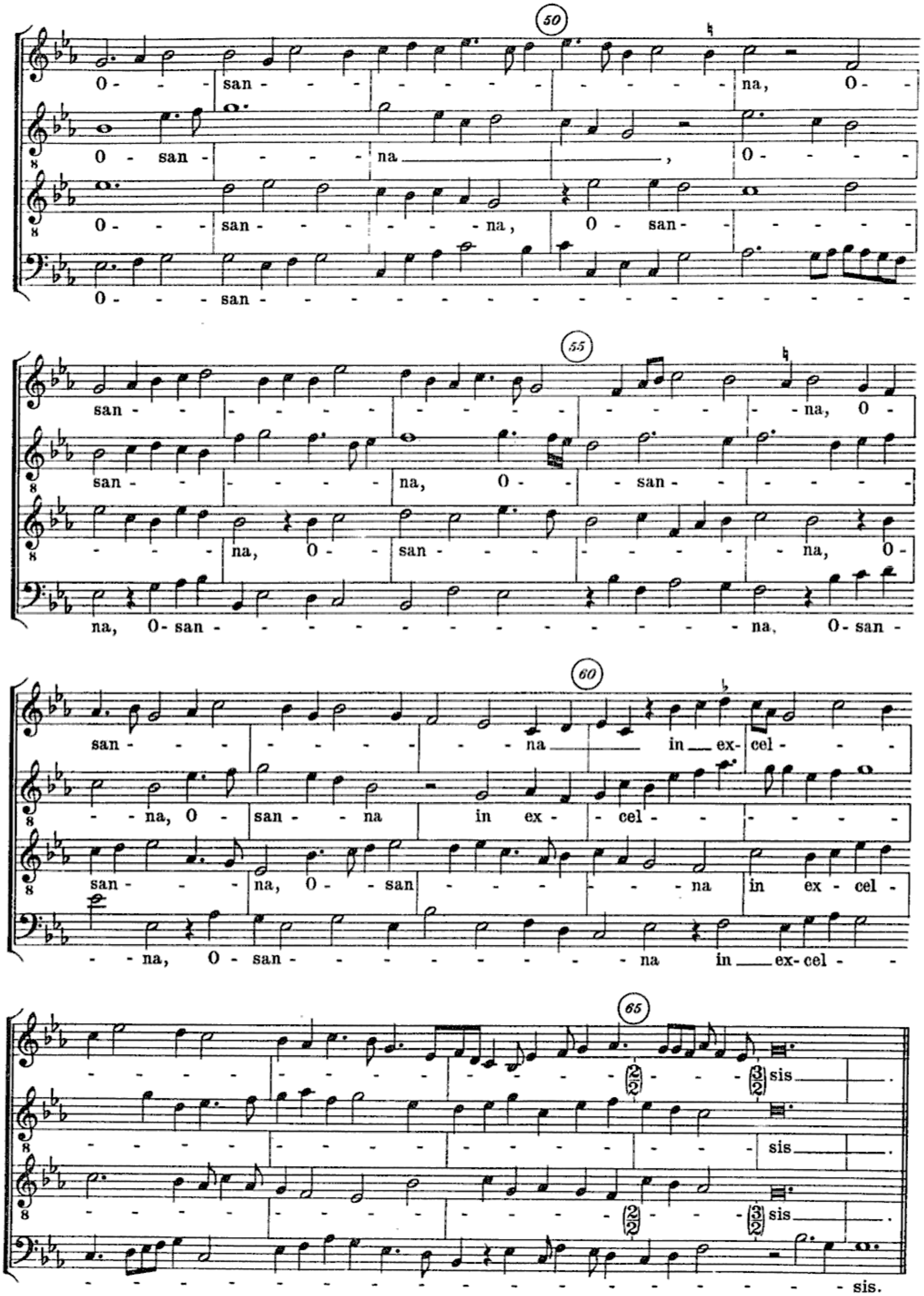

Figure 6. Johannes Ockeghem, Missa Mi-mi, "Sanctus," mm. 47–66.

Along with the resounding final cadence, the standout moment in the Sanctus is the transition that occurs between measures 46 and 47 (2'20" in the Clerk's Group recording), which marks the close of one section and the opening of the next. In measure 46 a cadence comes to a close on a perfect octave. The pause on the cadential arrival alters whatever attentional attitude the listener assumed while listening to the preceding measures. The movement and intertwining of the vocal lines that carried the listener along comes to a gentle stop. Whether the music was experienced as moving by while the listener stood back and observed, or whether it seemed to be carrying the listener along with it, the movement ceases.

Rolf Inge Godøy argues that such sensations of movement are examples of embodied perception, writing that "whatever we perceive or think is correlated with mental simulations of body movement."[141] He contends that we understand sounds by mentally rehearsing how we ourselves might perform the gestures they imply. This is based on an assumption, expressed by Marc Leman, that we perceive music "as an organism endowed with properties and as an acting subject involved in events."[142] Similarly, Arnie Cox asserts that our aesthetic interest begins with a question of "What's it like to do that?" and is answered by a mimetic response, actual or imagined, that attempts to replicate the movement we perceive; in the case of music, that mimetic response can extend to replicating the motions that must have been involved in making the sounds we hear.[143]

141 Rolf Inge Godøy, "Images of Sonic Objects," *Organised Sound* 15, no. 1 (April 2010): 55.
142 Marc Leman, "An Embodied Approach to Music Semantics," *Musicae Scientiae* 14, no. 1, suppl. (March 1, 2010): 51.
143 Arnie Cox, "Embodying Music: Principles of the Mimetic Hypothesis," *Music Theory Online* 17, no. 2 (July 1, 2011), http://www.mtosmt.org/issues/mto.11.17.2/mto.11.17.2.cox.html.

In addition to a mimetic response, there is another possible form of embodied music perception: an engagement that is founded upon perspective-taking. According to Sean Kelly, the gestures that are solicited from a perceiver are ones that are more revealing of the object of interest. Kelly writes how "some points of view are better than others," and that "better points of view immediately solicit me to take them up."[144] This is a phenomenon of listening as well as seeing. Combined with a background framing that delimits the range of the body's "I can," such gestures affect our own spatial situation.

In the Sanctus, there is a brief pause before the final section begins. The music has not finished—an audible breath captured on the recording provides confirmatory evidence that it is about to continue—but how it is to continue is uncertain. The marked contrast between this moment and the entrance of all voices in measure 47 creates a palpable spatial effect. Through a change in texture and harmony the music suddenly rises up with greater mass and strength to deliver the most harmonically rich music in the Sanctus. This change in depth is a consequence in the change in affordances, the textural contrast that occasions a new mode of listening, a new posture, an adjustment of our hearing apparatus arising with the shift in attention from one musical object to another.

At measure 47 a textural shift effects a change in figure-ground relations among the music's structural components. The homophony now stands out in relation to the polyphonic ground established in the preceding sections. Perceptually, several voices fuse into a single sound—a *Klang*—and we hear one sonic object. But more than that, the new texture changes the phenomenal sense of our relation to the environmental space in which

144 Sean Dorrance Kelly, "Seeing Things in Merleau-Ponty," in *The Cambridge Companion to Merleau-Ponty*, ed. Taylor Carman and Mark B.N. Hansen (New York: Cambridge University Press, 2005), 93.

we are situated. Rather than attending to an unfolding series of voices passing by—ephemeral in their coincidences, divergences, dissonances, consonances, and tessituras—we are confronted with a new object that has rapidly broken free from the close circle that enveloped us as we listened in to two voices (fostering a sense of intimacy akin to eavesdropping on a conversation). Like Miles Davis's organ clusters, this new, louder and more resonant object fills out the space more fully. But more than that, as this sonic object now seems larger than the intertwining voices we had just been attending to, the space we are in seems to expand in order to accommodate it. It therefore is reasonable to conclude that this moment holds such power because our bodies require a moment to adjust to this new sonic configuration. We need to determine how to best grasp this sudden surge in the music. This motivated search for perspective is an embodied activity, even if we don't find ourselves compelled to move our body at all. For standing still in order to take something in is just as much a perspective-taking activity as turning our head or body relative to an object of interest. To ensure a maximum grip on that which is presented to us means taking stock of the phenomenon and letting it affect us, and allowing it to present us with its own set of solicitations and affordances.

As we are affected by the sound, and as we are asked to regard the fullness of a triadic harmony, something remarkable happens. Rather than being shot through a sonic array as Eric Clarke describes in reference to a similarly sustained sonority in *Wozzeck*, we find stillness imposed on us by an emergent object, carrying its background along with it. The object carves out its space in relation to the background it itself determines. And the volume and resonant depth that suddenly confronts us, the physicality of the sound, *necessitates* a background upon which to comprehend it, a background, Merleau-Ponty

claims, that is always comprehended in terms of the range of possible bodily movement. We don't lean in as we might have in the preceding moments and focus on minute variations of consonance and dissonance that constitute two-voice counterpoint. We can't. The depth has expanded and along with it, the background that contains it. The possibilities for attentive focus, the affordances, are different. With this long, sustained triad, we can no longer take that same perspective of narrow focus.

It is not just that the music in measure 47 is louder (though that is a contributing factor), nor that additional voices, singing additional notes, create more reverberations that are picked up by the microphones, and that culminate in the aural effect of three-dimensional space via two speakers in the home listening environment. There is something more here. The new sound is one that suddenly enmeshes the listener in a more spacious environment. Regardless of the fact that the ground upon which this figure emerges is not easily definable, we know at the very least that it must be larger than the figure itself in order to contain it as a figure. This is a consequence of Morris's "logic" of envelopment cited above in chapter 2. If we sense a more voluminous sound here than those which we had just heard, and if both these sounds manage to fill the space in which we are situated (which, of course, most all sounds do), then this change in perspective results in the enlargement of our felt space. The spaciousness here is different than the spaciousness captured on the recording, that is, the spaciousness that is revealed through the reverberant characteristics of the performance space. It is space as we live in it, inhabiting space, as opposed to simply occupying it.

3.3 Distance and Objecthood

In the *Phenomenology of Perception*, Merleau-Ponty speaks often of the distance between perceiver and the thing perceived. He means this not as a metaphor, but neither as a measurable distance. He is noting the phenomenon of boundedness that separates perceived objects from oneself and from other objects, and the property of objecthood that emerges from these things within our purview.

In this view, when we sense there is something more in the music than the notes inscribed in a score—that a rich, sonorous chord, for example, is a single object that resonates around us in a different manner than a pair of independent, intertwined voices—we are not attributing an extramusical meaning to that object drawn from our personal history. Its appearance alone delineates more than its material components can account for. Merleau-Ponty writes of a line drawn on a piece of paper. It is not "a thing or an imitation of a thing. It is a certain disequilibrium contrived within the indifference of the white paper…. No longer the apparition of an entity upon a vacant background…. It is, as in modern geometries the restriction, segregation, or modulation of a pregiven spatiality."[145] The line is already, at minimum, an outline, indicating a figure, and alongside it, a background, continuing behind it uninterrupted. Music shapes space in a similar manner. To say we merely imagine a sound to get closer, or larger, or close in or open up, is to

145 Maurice Merleau-Ponty, "Eye and Mind," in *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, ed. Galen Johnson, trans. Michael Smith (Evanston, IL: Northwestern University Press, 1993), 144.

misunderstand these impressions as something illusory, rather than a manifestation of the same grasping action that allows us to perceive anything at all.

Merleau-Ponty notes that "the thing is constituted in the hold my body has upon it; it is ... a structure available for inspection by the body."[146] And as he describes it, the emergence of the object only occurs through the hold the object has on us: "When I simply glance at the surface that I am about to recognize as the surface of the table, it already invites me to a particular focus and calls forth the focusing movement that will give it the surface's 'true' appearance."[147] This idea of things beckoning the perceiver is common throughout *The Phenomenology of Perception*.[148] Here Merleau-Ponty is referring not simply to the focusing mechanism of the eye, but also more broadly to "focusing" in general: seeing the table entails looking at it, which perhaps requires movements of the head and neck in order to get it in focus. Perception is not passive reception, but the active search for information, and vision is not located in an organ, but is an achievement of the entire body. As Gibson notes, "*the visual system* is …. a hierarchy of organs and functions, the retina and its neurons, the eye with its muscles and adjustments, the dual eyes that move in the head, the head that turns on the shoulders, and the body that moves around the habitat."[149]

Merleau-Ponty, in his writing on aesthetics, claims that in perceiving the artwork, the viewer "takes up the gesture" of the painter. The "I can" of the body is enlisted by perception to make sense of the painting, to apprehend the meaning the work expresses.

146 Merleau-Ponty, *Phenomenology of Perception*, 334.

147 Merleau-Ponty, 331–32.

148 This is an idea shared by both Marratto and Kozak. "The movements of our bodies are a response to a call of the sensible," Marratto (2012), 8; "[music]… beckons for one's motor actions," Kozak (2015).

149 Gibson, *The Ecological Approach to Visual Perception*, 96.

We attend to what it affords, and by grasping what those affordances are, and sensing how best to gear our body toward those affordances, we comprehend an expression generated by the work.

Thus, with the triad, we can approach whatever affordances it offers. And what it affords is a standing still as we consider the contrast and "take in" the sound of harmonized "*Osanna*." And along with the temporary motionlessness imposed upon us, we may have a sense of being held aloft by the music; we hold still, knowing that moving away will not allow any better grip on this musical object. The rush of outward motion that the music manages to enact corresponds not with a body in motion but with a body suspended. As a result, if we aren't moving with the music, or following along with the movement, then it is the music that thrusts out toward us and holds us in its sway. If instead it were the paired contrapuntal voices that gradually became louder, with more complex rhythmic interplay, it is doubtful we would have the same sense of being amidst a spatial expansion of our surroundings. (In later chapters I will explore such textural and timbral changes and the phenomenal spatial experiences they engender.) To be sure, such a stillness is by no means passive. As Robert Hatten notes, even the held note designated by a fermata is understood as a pose, or a posture, revealing "the energy and affect with which it is invested."[150] To attune oneself to a new object is to assume a posture—to continually engage with the music in an embodied fashion.

Note also the congruence here between Besseler specifying different modes of engagement with music and Merleau-Ponty's idea that perceiving things as unified, inter-

150 Robert S. Hatten, *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert* (Bloomington, IN: Indiana University Press, 2004), 126.

sensory objects (i.e., as the same thing whether seen, heard, or felt) involves a "taking up" of the qualities they present:

> We understand the thing as we understand a new behavior, that is, not through an intellectual operation of subsumption, but rather by taking up for ourselves the mode of existence that the observable signs sketch out before us. A behavior outlines a certain manner of dealing with the world.[151]

For Merleau-Ponty, to see a thing is to take up the gesture it expresses. "The accomplished work is thus not the work which exists in itself like a thing, but the work which reaches its viewer and invites him to take up the gesture which created it and, skipping the intermediaries, to rejoin, without any guide other than a movement of the invented line."[152] As philosopher and phenomenologist Taylor Carman explains, "the work discloses a world for us precisely because it is a gesture we can identify with. We do not merely observe paintings, we visually participate in them."[153] This participation will become even more important as we proceed, as a work's affordances reveal how the solicitation to "take up the gesture" provides access to the meanings a musical work is able express.

3.4 Conclusions

To stand still, to heed the injunction to remain motionless, to consider for a moment the presence of an object in our midst, these are the affordances that allow Ockeghem's triadic harmony to provide its own variety of spatial experience. Besseler claimed that the triad is a separate object, with solidity and depth, but we can add that the triad here is an

151 Merleau-Ponty, Phenomenology of Perception, 333.
152 Merleau-Ponty, "Indirect Language," 88.
153 Taylor Carman, Merleau-Ponty (New York: Routledge, 2008), 186.

entity that then has the capacity to overflow its boundaries as a musical object and surround the listener as a sonic one. In a similar fashion, different musical works create different spatial relationships between themselves and their listeners. For this Sanctus from Ockeghem's mass, the space projected at its beginning is a narrow gap of the sort that opens up between interlocutors—intimate, yet just distant enough to allow nuances to be observed. Or rather, it is a spatial relationship that is afforded by a pair of soft melodic lines intermingling with one another. But with the homophonic sonority that begins the last section, space is deep. It is deep because it must contain both the voices of the singers and the listening subject. And it is deep because we are no longer paying attention in the same manner that we did to the single voices, but instead must contend with an object that spreads around us, demanding that we heed a different call than one that asks us to follow along. Ockeghem achieves this effect harmonically, rhythmically, texturally; the Clerks' Group's performance highlights the moment; and the listener sets the conditions for its emergence through arranging for an uninterrupted listening session at an appropriate volume setting.

Even without audio processing technology to affect the ear's directional hearing, a musical object can emerge that envelopes us. This is different from any directionality that may have been sensed regarding sound sources. It is a phenomenal experience of space and self and music and of relative stances, attitudes, and perspectives in which we attempt to grasp the import of the moment. It is the maneuvering we do as we assess the optimal means by which to grasp the music that accounts for Besseler's reaction of standing still. Yet we are not thrown outside the music, as Besseler would have it. Such a conclusion seems borne out of a predisposition to view music through a lens that bifurcates the

experience of music listening into two separate domains of appreciation: one deemed to be a cultured, and cultivated, response (combining, somewhat paradoxically, both Kantian disinterested contemplation and a Romantic encounter with the sublime) and a naive (or uncivilized, perhaps even barbaric) response that is unreflective and automatic, borne of an underutilized intellect prone to give in to the predilections of desire. On the contrary, the momentary daze we experience as we take in the beginning of the "Osanna" passage occasions a more fluid, less distant, engagement with the music—we need to open up to it so that we can consider what it affords.

Musical experiences such as these also alert us to the fact that the environments we inhabit are not immutable. We limit music's capacity for expressing spatial worlds when we limit its expressive power to a topographical acoustics, such as the voice that is panned from the left speaker to the right. Mimetic work is often performative, not simply representational, and can lie in the transformations themselves, in the transition from one space to another. These spatial transformations are rendered aesthetic and meaningful through the pleasure and play that comes along with the remaking of the everyday spaces we live in. It is the contrast Ockeghem's sustained triad offers that solicits an altered mode of attention, thereby changing the object of perception and reconfiguring our background frame of reference—that is, our surroundings themselves. It is through such active engagement with components of musical structure that the spatial aspect of the music reveals itself.

Chapter 4: George Crumb, *Black Angels*

4.1 Theoretical Framework

4.1.1 Introduction – Immersion as a Product of Timbre

In the following chapter, I will point out forms of engagement solicited by George Crumb's *Black Angels* through its bravura display of timbral variety and contrast. In particular, I will present an analysis of the sense of immersion, how it differs from the sense of being surrounded, and how its dislocating effect further illustrates the spatial qualities evoked through the listener's perspective. I will refer to a specific recording of the work (by the Kronos Quartet, 1990) to describe how timbre, in the form of reverberation, situates a listener with respect to musical sound.[154] To do this, I will draw once again on Merleau-Ponty's theory of embodied perception, particularly on his discussion of depth and the adjustments made by our bodies to turn what is at most "a vague beckoning" into an "actual" object or "present reality."[155] I will then show how this relation becomes complex when a listener is engaged with a piece of music featuring instrumental sounds with certain distinctive timbres that invite the listener to reconsider the spatial relations in which they are currently enmeshed. In particular, sounds with inharmonic overtones, sounds lacking a clear attack, and sounds whose direct vibrations can cover up the reflective vibrations of

154 George Crumb, *Black Angels*, Kronos Quartet, Elektra Nonesuch 9 79242-2, 1990, CD.

155 Here, I rely on Colin Smith's translation of "*sollicitation*" as "beckoning," rather than the cognate "solicitation" found in the Landes translation of *Phenomenology of Perception*. The former term links Merleau-Ponty's "*sollicitation*" with the idea of music as a call that I have been developing in this study. Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. Colin Smith (New York: Routledge, 2002), 248. The "actual" object and "present reality" are Landes's and Smith's translations, respectively, of Merleau-Ponty: "*on corps se centre sur un objet encore virtuel et dispose ses surfaces sensibles de manière à le rendre actuel*." Maurice Merleau-Ponty, *Phénoménologie de La Perception* (Paris: Gallimard, 1945), 276. Merleau-Ponty, *Phenomenology of Perception*, 2012, 249. Merleau-Ponty, *Phenomenology of Perception*, 2002, 278.

other sounds. I show how these musical sounds can directly affect a listener's sense of space by altering cues we rely on to determine the volume of our surroundings, thereby altering our understanding of the spaces we inhabit.

As Merleau-Ponty describes with vision, depth is experienced along with, rather than abstracted from, our perceptions. Though the apparent size of an object varies with distance, we experience that size as constant, and the distance as changing. We don't calculate its size relative to objects around it and then deduce its distance from ourselves. As Komarine Romdenh-Romluc explains, "[Merleau-Ponty] holds that the phenomenon of apparent size does not merely indicate depth; it is an experience of depth."[156]

Furthermore, we must find ourselves placed in the same world as the objects themselves for them to appear with any depth. Merleau-Ponty gives the example of a painting. We cannot perceive depth in the scene the painting depicts because its space does not extend out toward us, and continuing both behind and beyond where we are.[157] In such a case, we are relegated to outside observation. Romdenh-Romluc explains how Merleau-Ponty considers apparent size unmeasurable, that it is not in fact an experience of actual size, but an experience of "size-seen-at-a-distance."

The recording environment is always a component of the timbre of music as heard. In the following analysis, I note how "solicitation" emerges as a type of call issued by the music by noting contrasts in framing techniques, performance choices, and, most clearly, in modes of address, as in the spoken word portions of the piece. The manner in which the

156 Komarine Romdenh-Romluc, *Routledge Philosophy Guidebook to Merleau-Ponty and Phenomenology of Perception* (New York: Routledge, 2011), 110.
157 Romdenh-Romluc, 109.

music addresses us affects how we respond, which in turn affects what we perceive. Space is thus revealed as an affective dimension of our environment.

This affective dimension is important in considerations of music and meaning. As I showed in the previous chapter, embodied cognition is implicated in how we understand not only the music we hear, but also the space in which we hear it. I will expand on this point with Merleau-Ponty's remarks on grasping as a means to make sense of our world. Different objects we encounter solicit different modes of engagement, each with a corresponding bodily gesture necessary to grasp and make sense of them. We "grapple" with things we encounter in order to make them show up with maximum clarity. Such grappling is indeed a process of the body, of making adjustments to better tune into and attend to those things that we deem to warrant our interest. I explain how this grappling is always at work by presenting audiological and acoustic research on the role of movement in perception. With this understanding of embodiment as an ongoing process, taken in response to our own affective concerns, I then correlate specific features of the sound of Crumb's music, its timbral components, with the movements they elicit and the spatial perceptions that result. This spatial understanding of the proprioceptive component of embodied perception allows us to explain in more concrete terms the spatial effects discussed in previous chapters. As it is one of the more distinctive effects of music, and one that is closely linked to timbre, yet not often discussed, I begin with a discussion of the sensation of being immersed in music.

To be immersed in music is different from being surrounded by it. At first, it may seem that to be immersed and to be surrounded by music are essentially the same thing. The words are often used interchangeably. In this chapter, I want to point out a subtle

distinction between the two. When used to describe audio systems that feature "surround sound," the word reflects the container-concept of space described in Part I, a concept I have been critiquing due to its limited ability to convey aspects of spatiality we experience when listening to music. It is a word that describes objects and their locations relative to a subject. (NB: This differs from the term surroundings, which is a term for the collection of things that surrounds a person or organism as a state of affairs or condition, generally used in the sense of an environment.) To be surrounded by music, then, is to hear it directed at you from all sides. And as is typical with surround sound systems, or with spatial music, those sound sources must be perceived as distinct from one another. (For example, sound bouncing off the walls of a performance space, subordinate to a perceived direct sound source, is not an instance of listeners being surrounded by sound in this sense.)

Immersion, on the other hand, is a felt quality. The word speaks to the ability that sound, specifically acousmatic sound, has to envelop the listener. Not simply surrounded now, but enclosed within it. What contributes to this sensation of being within the music? It is more than simply the result of being placed among the musicians, or among strategically placed loudspeakers. Immersion, as we shall see, arises precisely when those locative indicators of space are removed.

Timbre is a particularly effective means of creating spatial impressions, and is implicated in both experiences of being surrounded by and being immersed in music. As I will explain below, timbre is responsible for providing us with crucial information regarding the distance of sounding objects. Our physical location has a much bigger impact on the timbres we hear than on pitches, harmonies, and rhythms. While sound can be all around us through its means of dispersal, timbral distinctions help us place ourselves

within that environment, through directing our attention, soliciting optimal perspectives, and affording affective responses. In this analysis of the second movement of George Crumb's *Black Angels*, I will discuss the sound of the tam-tam and its ability to pervade an environment so fully that the timbral features we rely on for localization are submerged and, in a sense, so is the listener. But I will discuss other experiences as well, when the sonic environments created through the music are shattered open through timbral shifts, and sonic worlds are revealed to be not as fully enveloping as they may at first appear.

While either headphone or loudspeaker listening may produce immersive experiences, loudspeakers allow for a unique modulation of the listening environment, that is, the merging of the sonic characteristics of distinct environments described in the analysis of "Maiysha" in Part I. Since a change in the body's position does not alter the sound that reaches the ears in the headphone scenario, the aural spatial configuration is fixed. Moreover, this spatial configuration is highly distorted. Audio researcher Bosun Xie writes that headphones can only offer lateralization—a scheme where the sound field is restricted to a horizontal, left-right axis. Yet, strangely, it is a horizontal axis that is situated within the listener's head:

> What occurs in headphone listening is not localization but "lateralization." Localization refers to determining the position of an auditory event in three-dimensional space. By contrast, lateralization pertains to determining the lateral displacement of an auditory event in one-dimensional space, that is, along the straight line connecting the entrances of the two ear canals.[158]

158 Psychoacoustician Jens Blauert adds that "localization of auditory events inside the head often occurs as an undesired side effect when headphones are used to present signals." Bosun Xie, *Head-Related Transfer Function and Virtual Auditory Display* (Plantation, Florida: J. Ross Publishing, 2013), 8. See also Moylan, *Understanding and Crafting the Mix*, 2007, 305. Jens Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (Cambridge, Mass: MIT Press, 1997), 117.

Because of this distinctive phenomenon associated with headphone listening, my inquiry focuses on the practice of listening to recorded music through loudspeakers, a mode of interaction in which music performed at another time and place is made to resonate throughout the listener's personal surroundings. Limiting the study in this way nonetheless presents a challenge in defining the perceived spatial effects of music. Even audio engineers struggle with the concept of space when considered in terms other than locations in a three-dimensional coordinate system. The following passage by technician Frances Rumsey illustrates this point:

> Envelopment, spaciousness, and spatial impression are terms that seem to result in the most varied interpretation in the literature. They are harder to conceive of than dimensional quantities such as width, depth, or distance, as they are not perceived directly as linear quantities but more as semiabstract and multidimensional impressions. They are all terms that relate in some way to the degree of immersion in the sound field experienced by the listener or to a global description of the scene.[159]

This chapter will describe some of these hard to define experiences of space in relation to theories of embodied meaning and Mark Hansen's understanding of virtual reality (VR) as mixed reality, as described in Part I. The reality of our own environment, in his view, is equal to the reality of those that are digitally simulated, technically reproduced, or socially constructed. We enter and explore all of these with our bodies, and make use of the body's ability to unify these separate realms into one reality through movement.[160]

159 Francis Rumsey, "Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm," *Journal of the Audio Engineering Society* 50, no. 9 (September 15, 2002): 661.
160 Hansen, *Bodies in Code*, 2005, 8–9.

4.1.2 Immersive Properties of Sounding Music

One commonly identified trait of acoustic immersion is the sense of perceiving something from the inside. Will Scrimshaw, in his book *Immanence and Immersion* on the acoustic component of contemporary art, states that in an immersive installation, "listeners or viewers assume a constitutive role in the production of the artwork, rather than observing an autonomous object from a distance…occupied and explored from within rather than observed from without."[161] But immersion entails more than just change of perspective. When immersed, we cannot locate ourselves at the center of any sonic space in particular, because the spatial cues for doing so are not available. This is because in this timbrally-composed variety of immersion I seek to describe, the auditory information we use to determine left and right, up and down, forward and back, is unavailable, no matter how we move. And the environmental information we glean from reflected sound—information regarding the size and material construction of the environment—is merged into the instruments' direct sound such that reflections no longer register as such. As opposed to the surrounded listener, occupying a distinctive location, the immersed listener is left adrift. The sense of space offered by immersion is therefore much different than the one provided by a surround-sound setup: it requires our body to recalibrate to a situation in which its position relative to its surroundings is no longer secure.

According to Amy Cimini, Merleau-Ponty's contention that we are surrounded on all sides by a world means that we are situated in a world before we are even able to begin to

161 Will Schrimshaw, *Immanence and Immersion: On the Acoustic Condition in Contemporary Art* (New York: Bloomsbury Academic, 2017), 89. Eidsheim conceives of immersion this way as well: "The result is a feeling of being immersed in a sound (rather than observing it from the outside or at a distance)." Nina Sun Eidsheim, *Sensing Sound: Singing & Listening as Vibrational Practice* (Durham: Duke University Press, 2015), 68.

grasp the significance of those objects of interest we encounter in it. Immersive sound, by obscuring the spatial traces of the environment that sound ordinarily conveys, can therefore draw attention to our inherence in the world: it compels us to reestablish our equilibrium while simultaneously frustrating our ability to do so.[162] Musically, immersion can invite us into the music in a different manner than, say, highly infectious rhythms. While rhythms engage the active, participatory quality of our body in exploring our environment, immersion can draw us in by blurring the boundaries between ourselves and our surroundings. As I will show in the Crumb analysis, instruments such as tam-tams, gongs, and cymbals can be used to foster this sensation of immersion. Beyond their well-worn martial and ceremonial associations, these percussion instruments can be used to reach out to listeners, drawing our attention outward from the local to the global.

Yet immersive sound presents a problem for Merleau-Ponty's embodied theory of meaning described in the previous chapter. According to Merleau-Ponty, we make meaning out of "hollows," "furrows," "lines"—the "disequilibrium" that such twists introduce into our surroundings. Unlike immersion in, say, water (or air, as it is proposed in aesthetic theories of atmosphere described in the next chapter), with sonic immersion such disequilibria are barely noticeable because our aural environment uniformly extends around us in all directions. Because there is nothing to grasp, we are plunged into a featureless morass and a passivity is thrust upon us due to the ineffectuality of our actions. Our gestures no longer seem to be the means by which we make the object appear for us: it can only confirm our inability to achieve *any* perspective that might be deemed more optimal than any other. The foreground-background distinction seems to dissolve, with the

[162] See Part I, p. 63: Taylor Carman on Merleau-Ponty and "optimal grip."

result that we perceive sound not as something arriving at us from a source, but as an atmosphere we inhabit. And here we reach the feeling—described earlier by Scrimshaw—of being inside the thing, and a change in the thing from an autonomous object to something less material, yet still eminently sensible.

Throughout the present study, there has been a tacit understanding of the fundamentally interactive nature of all music listening, and indeed, of all perception. All the music analyses thus far have featured engaged listeners actively striving to comprehend musical works—works that, in turn, respond to their changes in perspective. By taking an embodied approach, we hew to the proposition that we perceive things as "eliciting the movements of our bodies," and *only* through their eliciting the movement of our bodies.[163] This is the stance taken by Gibson as well, with his notion of affordances and how we perceive things according to what it is we need or want to do. And yet, immersion can frustrate the interactive component of listening, by inhibiting the ability to gain a different perspective that would reveal a different aspect of the music. As the VR artist Char Davies observes in regard to a loss of purchase on our surroundings—an inability to obtain a grip on things—associated with immersion: "Always what was important to me was the notion of being immersed in enveloping space, and the sensation that you're fully enveloped…*it's not about interactivity* but the fact that you are spatially encompassed and spatially surrounded—it's all around" (emphasis added).[164]

Immersion, then, will serve as a limit case for this embodied interpretation of perception and meaning. And our investigation of musical timbre will show that in spite of

163 Marratto, *The Intercorporeal Self*, 32.

164 Quoted in Frances Dyson, *Sounding New Media: Immersion and Embodiment in the Arts and Culture* (Berkeley: University of California Press, 2009), 107.

their subsuming character, we are nonetheless capable of making sense of these immersive moments, and further, of grasping musical meaning in an embodied manner. For it is only through the presentation of "incompossible" spaces (to use a term philosopher Scott Marratto takes up from Merleau-Ponty's *The Visible and the Invisible*) that the work of perceiving and understanding is accomplished. This constant striving of the body to establish equilibrium is called upon when grappling with the competing claims of reality that arise when confronted with a sonic, spatial world presented by the work and a surrounding spatial environment inhabited by a listener—an interpenetration of worlds that is enacted in every instance of listening to recorded music. An explanation for this sense-making ability lies in an account of meaning that consists in grasping together these incompossible affordances and incorporating them into our own lived experience of the environment.

The sense of futility that immersion imparts is different from the passive feeling of being acted upon, and is different from the momentary motionlessness solicited by Ockeghem's mass. With the triad in the Ockeghem example, we do have a musical object that is presented to us for our perusal. However voluminous this triad, there is something to grasp and a sense in our bodies of how best to adjust in order to grasp it—even if we can never hope to take it all in. By contrast, immersive sounds, or sounds experienced as immersive, don't set us off in search of objects. Later in my analysis I will focus on the prolonged peal of a tam-tam, but there are other non-pitched instruments that produce similar immersive sounds through chaotic modes of vibration, such as the cymbal and gong, and even the diaphanous and vaguely pitched whine produced by these instruments when excited by a bow.

4.1.3 Hearing and Localization

William Moylan states that the identification of sound source location is achieved binaurally: a travelling sound wave will strike left and right ears at different times—with different amplitudes.[165] When a sound is present that is not directly in front of or behind a listener, the sound will reach each ear at a different point in its wave cycle, resulting in a phase difference which the human auditory system is able to detect. This is known as Interaural Time Difference, or ITD. It functions in tandem with the Interaural Level Difference, or ILD, which denotes the variation in amplitude of a sound as it reaches each ear—the head itself blunts the intensity of waves that have to travel around it.

However, as further auditory research has shown, the physiology of sound location extends beyond the placement of ears on opposite sides of the head. Our entire bodies are involved, and audio technologists have been occupied with finding models for replicating sounds based on how they are heard by individual listeners with different body shapes. The Head-Related Impulse Response (HRIR) is a formula that predicts how sound waves are altered when filtered by the listener's own anatomy—reaching the ears after resonating through oral and nasal cavities, being absorbed by the torso, refracted by the head, bounced up from the abdomen, and reflected off the complex folds of the pinnae. No two bodies are exactly alike; consequently, no sound reproduction system will ever be able to reproduce how any recorded sound would be heard by any particular listener.[166] Thus, there is an openness to the interpretation of any type of reproduced sound.

165 Moylan, *Understanding and Crafting the Mix*, 2007, 3–35.

166 Although HRIR databases have been compiled to assist in approximating these effects.

Another component of sound localization is distance, an attribute that Moylan identifies squarely with timbre. Debunking the commonly held notion that we judge the distance of sounding objects we hear according to their loudness, Moylan (using capitals for emphasis) states that "distance is NOT loudness."[167] Instead, he explains that the perception of a sound's distance is a function of that sound's timbral detail. Higher frequencies, having less amplitude, do not carry as far as lower ones. In addition, the proportion of direct to reflected sound can also carry cues for distance. A closer sound will consist of a higher proportion of direct to reflected sound, with more of its original high frequency components, while a sound that has to travel a longer distance will lose some of its high-end sheen. Higher frequencies require higher energy levels for dispersal, and thus decay more rapidly than lower frequencies. These facts are important to know for any analysis of music perception and are often aspects of sound we take for granted. But the behavior of partials in situations that involve the playback of recorded music involves a further layer of complexity: the timbres specified in the score are inscribed in the recording along with the location cues of the performance setting, and this particular sonic representation of the work is then realized in yet another spatial context—the playback environment—with its own acoustic properties.

Joseph Klett proposes the analytic concept of a "sonic object setting" when describing this inseparability of a sound from its acoustic interactions with the environment in which it is sounded. He writes: "what we hear as a sonic object is the 'loose coupling' of causal source and perceptual event."[168] In addition, he notes how such hearings

167 Moylan, *Understanding and Crafting the Mix*, 2007, 189.

168 Joseph Klett, "Sound on Sound: Situating Interaction in Sonic Object Settings," *Sociological Theory* 32, no. 2 (2014): 149.

are determined situationally, and involve social and cultural context along with the material environment. His example is the sound made when a car drives over a scrap of metal lying in the street. What the driver hears is different from what someone on the street hears, which points to the independence of distance and the loudness and timbre of sounds. Depending upon the concerns of the hearer, the sound is not only an indication of the scrap's metallic property, but it can also signal road conditions—how safe it is for a pedestrian to cross due to the number of cars and how fast they are going, In addition, Klett notes the social construction of sounds: how, for example, a particular sound can be construed as the sound of a piece of metal, or the sound of a car driving down a debris-strewn street, or the sound of a derelict neighborhood. Different contextual frames provide different sound sources. "Perceptions of sound in sonic object settings may vary based on actors' presence, orientation, and attention."[169] And, in accordance with Merleau-Ponty's theory of perception, Klett describes the setting as a milieu of possibility: "The consequence of this spatially bounded audibility are the possible interactions afforded within particular situations. Setting delimits the scope of the situation, with more dynamic settings providing more situational variation."[170]

The coupling, then, between a sound and its source is neither fixed nor immediate. Any cause-and-effect relationship between them is contingent upon the situation in which the event occurs. And it is this characteristic of aural perception that allows for sonic modulations of space to occur. Because of the contingency of the relation between sound and source, some musical sounds, such as the peal of the tam-tam, are able to deliver

169 Klett, 150.
170 Klett, 150.

experiences where not simply the source of the sound is indefinite, but also the characteristics of the space in which it is sounded. When timbral cues to the setting are distorted or obscured by sound, we can easily lose our footing. By slackening the already "loose coupling" between sound and setting, such musical sounds can present a listener with a sense of disorientation—where we are situated in relation to the sound can become confused.[171] As I will discuss further in my analysis, the sense of immersion arises when we no longer have access to audible spatial cues of any sort, a situation that is generated by the tam-tam and its loud, inharmonic vibrations.

A sonic object setting is the conflux of the material arrangements that allow sonic qualities to emerge through our interactions with sound, and the situational frame that affects what we attend to. Klett's description is an elaboration of Moylan's strictly acoustic concept of a host environment (introduced in Part I) viewed through the lens of cultural anthropology. Klett's formulation applies to my description of the mixed reality environment that obtains when listening to recorded music. Klett notes how a more varied situation affords more possibility for interaction, just as, when the sound of a recording overlays the sound of the listening space, we are given multiple channels of information to access through enaction. Our own embodied situation shapes what we hear.

4.1.4 Speech

We have just considered how our embodied situation can affect the timbres that we hear. But the timbres we hear similarly affect our situation—our "orientation and

171 Klett cites Stephen Handel (1995) as his source for the collocation "loose coupling." Klett, 149.

attention" can switch through a change in character of the sound object and its setting. For example, when released from the task of following a melody, we are often afforded a chance to contemplate timbres and their subtle shifts in color. Such is the case with the tape loops of William Basinski, or with a piece like *Four Organs* by Steve Reich, where phase differences between identical instruments manifest themselves as sonic partials. One of the ways in which music makes us aware of space is through timbral contrast. Timbres can highlight the inherent spatiality of music by disorienting the listener, providing cues that solicit a rapid recontextualization of the listening event. As an example, recall from Part I how the situation of listening at home is framed as both a magic trick made possible through modern technology, and as a direct and unmediated encounter with the music, free of the contingencies of live performance. We may be conditioned to listen as auditors, or as witnesses, or as consumers, or as subjects—pupils of a sort, to be edified by the music. These roles we assume are also a result of the contextual frame constructed within the listening event. When the music presents a challenge to one of these roles, the effect is disorienting because our situation has changed and, as a result, our perspective is suddenly thrown out of focus. It is a change in situation that motivates a search for a new optimal perspective.

One way in which our carefully worked out equilibrium can be dispelled is through the introduction of spoken word into a musical performance—the utterance becomes immediate and the foregoing music is perceived as not only distant, but also no longer in the same contextual frame. The music's artifice—its identity as a crafted performance work—becomes apparent in comparison to the personal address of the spoken word. Such a moment is instructive. As with the phenomenon of immersion, the sudden change of

mode of expression reveals how music can overflow its frame and become a formative component of a listener's own peripersonal space. Here, reality becomes even more complex as a new component is added to the already "mixed" reality. There are now two contrasting situations that require integration in order for a listening frame to be re-established. Merleau-Ponty illustrates this point with a musical example, but his contrasting spaces emerge from the non-alignment of hearing and seeing, as when one is in a concert hall and sensing the space differently with eyes closed than with eyes open.[172] What I am describing is a conflict in worlds that arises from the radical reframing of the performance. This is significant in our discussion of space, because moments like these, according to Merleau-Ponty, are responsible for creating depth—what he also terms a "situational spatiality."[173] Depth was explored in Part I and can be briefly said to refer to the fact that the objects we perceive appear to us as either near or far, rather than simply arrayed alongside one another. The significance of this understanding of depth lies in the fact that it implies a situated viewpoint, a perspective from which things can appear near or far. Because the "'situation sensitive' character of depth-perception opens it to developmental and historical variations," as Marratto notes, we can also speak about the culturally and socially engrained practices that inform our listening as contributing to our sense of space.[174] This allows us to speak of the contextual framing of the listening event as an additional space-generating act. The sense of depth, of things being close or far, is not simply a physical or physiological phenomenon, but is culturally mediated, as well.

172 Merleau-Ponty, *Phenomenology of Perception*, 2012, 230–31.
173 Merleau-Ponty, 102.
174 Marratto, *The Intercorporeal Self*, 36.

Consider a musical example. When someone is singing, especially in a performance, there is a distancing effect that occurs with respect to the text. It is difficult to listen to any song as direct speech—the sung delivery of the words places the text in another category of expression altogether. It is a performance, and all the attendant qualities of performance that distinguish it from everyday activity lend it a distance from the listener. Even from the front row of a small performance space, a singer rarely appears in our peripersonal space in the same way as does a fellow-concert goer located a couple of seats away. The performer and the audience member may be equidistant from us, but what are the listener's possibilities for action? Whom could she tap on the shoulder to ask the time?

Similar differences in the affective character of the voice can be observed when words are chanted, whispered, mumbled, or shouted. These are all marked categories of speech that carry an extra level of meaning along with their semantic content. That extra meaning is based on myriad contextual cues that listeners need to take into consideration when interpreting the utterance. Crumb's piece takes advantage of this distinction between marked and unmarked forms of speech. His use of spoken language can not only surprise listeners, but can affectively take hold of them in ways that a more musical use of language cannot. Personal address reaches us along with its attendant cultural apparatus—the norms that facilitate dialog. Thus, through a sudden switch in frame—from the distanced effect of a musical performance to the closeness of conversation—a performance space can be transformed from grandiose to intimate (a move in the opposite direction, accomplished by different means, from that taken in Ockeghem's Sanctus). But beyond the sense of size of the enclosure, the imposition of a new contextual frame radically alters the listener's situation: to be engaged in conversation is a different activity than being engaged in music

listening. Even if we still sense a slight echo, space collapses down to the peripersonal, not through an alteration in the sonic object space projected by the recording, but by the unexpected reconfiguration of the interactional frame.

4.1.5 Meaning, Grasping, and the Role of Timbre in Creating Space

Meaning in the embodied sense encompasses more than simply propositions that can be expressed in language. For example, how a musical work makes us feel is a significant component of a work's meaning, a meaning different than the sort offered by a declarative sentence ("the cat is on the mat"). How then might such non-verbal meanings, such as feeling immersed, be apprehended by a listener? There are many accounts of aesthetic meaning that consider it to be the result of an act involving the grasping together of disparate elements performed by an engaged beholder. For Merleau-Ponty, this grasping is gestural and physical. The grasp is how we perceive, and what we perceive already has meaning in our grasping for it. To illustrate, Merleau-Ponty writes about a hypothetical experiment where a subject "only sees the room he is in through the intermediary of a mirror reflecting the room at a 45-degree angle from the vertical."

> At first, the subject is not geared to the utensils [the room] contains, he does not inhabit the room…. After several minutes…rather than his genuine legs and arms, he feels the legs and arms required for walking and acting in the reflected room—he inhabits the spectacle. And this is when the spatial level shifts and is established in its new position.[175]

175 Merleau-Ponty, *Phenomenology of Perception*, 2012, 261–62.

Meaning thus conceived is an activity of meshing with the world though reaching out towards it, and this grasping for purchase on what we perceive can only be done within a spatial realm that allows for such activity. This conception of meaning has been advanced by other philosophers as well. Hans-Georg Gadamer speaks about a fusion of horizons that occurs when a new context is taken up that can accommodate the unfamiliar.[176] Paul Ricoeur builds on Gadamer in his theory of metaphor, which posits a reader who is "consequently enlarged in his capacity of self-projection" through taking in a new meaning presented by a text.[177] For Mikhail Bakhtin, the content of a work lies in "the concrete individual participating in a concrete event, and aesthetic activity consists of bestowing upon this content a form which organizes it architectonically into a meaningful whole."[178] Marratto claims that even consciousness itself is a "holding-together" that depends upon "a keeping-apart that it can never succeed in overcoming."[179] While the concept of grasping something by the intellect may be based upon an earlier metaphorical use, the grasping that comes with embodied engagement with sound and music is not. Acoustic engineer Bosun Xie also notes the active role movement plays in localizing sound sources: how we aim the ear so that the pinnae can lay hold on the sound.[180] Arnie Cox has noticed this pervasive use of the word grasp to describe an act of understanding. The sense of grasp, for him, is concerned with the music-as-object metaphor (illustrated by his example of the

176 Hans-Georg Gadamer, *Philosophical Hermeneutics* (Berkeley: University of California Press, 2004).

177 Paul Ricoeur, "Metaphor and the Main Problem of Hermeneutics," *New Literary History* 6, no. 1 (1974): 107, https://doi.org/10.2307/468343.

178 Mikhail Bakhtin, *Toward a Philosophy of the Act* (Austin: University of Texas Press, 1993), 64. In Liisa Steinby, "Hermann Cohen and Bakhtin's Early Aesthetics," *Studies in East European Thought* 63, no. 3 (2011): 234.

179 Marratto, *The Intercorporeal Self*, 115.

180 Xie, *Head-Related Transfer Function and Virtual Auditory Display*, 17.

common expression, "it went right over my head"), one which is logically prior to an understanding-as-grasping metaphor.

What I am proposing, though, and the sense that comes through in the work of these phenomenologists, is a grasping together not of objects, but of situations—situations that present us with an incompossibility that we are called upon to mediate. It is a mediation that necessarily opens up a space for a new meaning to emerge. For example, in the spaces described by Hansen in Part I, the mix of different realities are integrated through the movement of the body.

Grasping, as we know from Gibson as well as Merleau-Ponty, is an affective process. The unfamiliarity of the new occasions an opportunity to discover new perspectives, a search that is, according to Merleau-Ponty, in itself a form of expression. By affording opportunities for grasping new meanings, by soliciting movement in terms of possibilities for action, music allows us an opportunity to fit together one world of possibilities with another. We do not merely recognize a virtual world—we allow our own world to be reshaped.

In discussing the meaning of music, Mark Johnson notes that:

> The reason that so many philosophers are unwilling to call these embodied structures part of meaning is that images, image schemas, affect contours, and metaphors cannot be satisfactorily put in propositional form. However, instead of concluding that music must therefore not have meaning in the proper sense, we ought rather to conclude that meaning, in the proper sense, goes far beyond conceptual and propositional content.[181]

181 Mark Johnson, *The Meaning of the Body: Aesthetics of Human Understanding* (Chicago: University of Chicago Press, 2007), 262.

The meaning we derive from works that are able to affect our relationship with our surroundings is a meaning that affects our mode of engagement with the world. We understand music's meaning when we comprehend the appropriate mode of engagement that will give us the work in an optimum balance of richness and clarity.[182] When we listen, our subtle movements and motions allow us to grasp the qualities of the sound that reaches our ears (and that moves through our head and torso). Merleau-Ponty's description of the role movement plays in visual apprehension is therefore especially applicable to a movement-oriented description of aural understanding:

> Occasionally a new knot of significations is formed: our previous movements are integrated into a new motor entity, the first visual givens are integrated into a new sensorial entity, and our natural powers suddenly merge with a richer signification … whose advent suddenly reorganizes our equilibrium.[183]

This enrichment of our natural powers helps to explain the visceral quality of musical meaning. When spatial characteristics of a recorded environment are laid atop our own, we are afforded new possibilities for movement as we work toward grasping the meaning presented to us, which results in a new relation to our surroundings. The proprioceptive awareness that, as Nelson reminds us, binds together body and space becomes a productive force in generating the space we find ourselves in. Each gesture that is undertaken in order to understand our surroundings gives form to the space in which it occurs. When our movements are tiny, perhaps even subliminal (for example, cocking the head, momentarily stopping the breath), we may not notice the impact these actions have on the way we perceive the space we are in. But such filtered attunement is what brings

182 Merleau-Ponty, *Phenomenology of Perception*, 2012, 332.
183 Merleau-Ponty, 155.

certain perspectives into being, and those perspectives we take on the sound are perspectives that are taken upon the space.

4.2 Analysis

4.2.1 Introduction

George Crumb's piece for electronic string quartet, *Black Angels* (1970), provides an opportunity to examine how a listener's spatial experience can be shaped through instrumental timbres, as well as and how those timbres are shaped by the contingencies of performance, the artifacts of the recording process, and the listening environment. In particular, the contrasts displayed in this work offer a springboard into a discussion of immersion and how such an experience can be facilitated by the sonic elements of a composition. The recording of *Black Angels* by the Kronos Quartet serves as a particularly good case study. It makes prominent aesthetic use of topographical space through its choice of microphone placement and representation of the host environment in a stereophonic recording.[184] In addition, the large, empty hall indicated by the reverberations of the musical sounds provides a clearly defined background setting—a frame—through which we can interpret the spatial contributions of the piece's timbral elements. This background, however, is often disrupted by the timbral elements themselves, and leads to some of the works most numinous moments.

The work is scored for an electronically amplified string quartet. The musicians are also required to play several different percussion instruments. Besides the use of extended

[184] George Crumb, *Black Angels*, Kronos Quartet, Elektra Nonesuch 9 79242-2, 1990, CD.

techniques, Crumb also gives instructions to the players to speak, whisper, shout, and chant. The texture of the piece consists of a wide variety of instrumental colors arranged into different streams of activity, some blended together and others willfully disjunct. As a result of changes in timbre, the relationship between those musical elements that are foregrounded and those which comprise the background can become unsettled—as can the contextual frame itself—which can in turn open up novel spatial experiences for the listener. In this analysis I focus on the central portion of this work: movement II, "Absence," which contains the climactic "Threnody: Black Angels" section. While all three movements contain a wide variety of timbral qualities, movement II contains stark dynamic and timbral contrasts that facilitate comparison between different means of spatial modulation.

"Absence" is divided into four sections, or "images" in Crumb's parlance (the subtitle of *Black Angels* is *Thirteen Images from the Dark Land*. See Figure 7.) The first and third sections of the movement (images 6 and 8) feature triadic harmonies and staid rhythms, while the second and fourth (images 7 and 9) emphasize unconventional timbres and unexpected series of musical gestures.

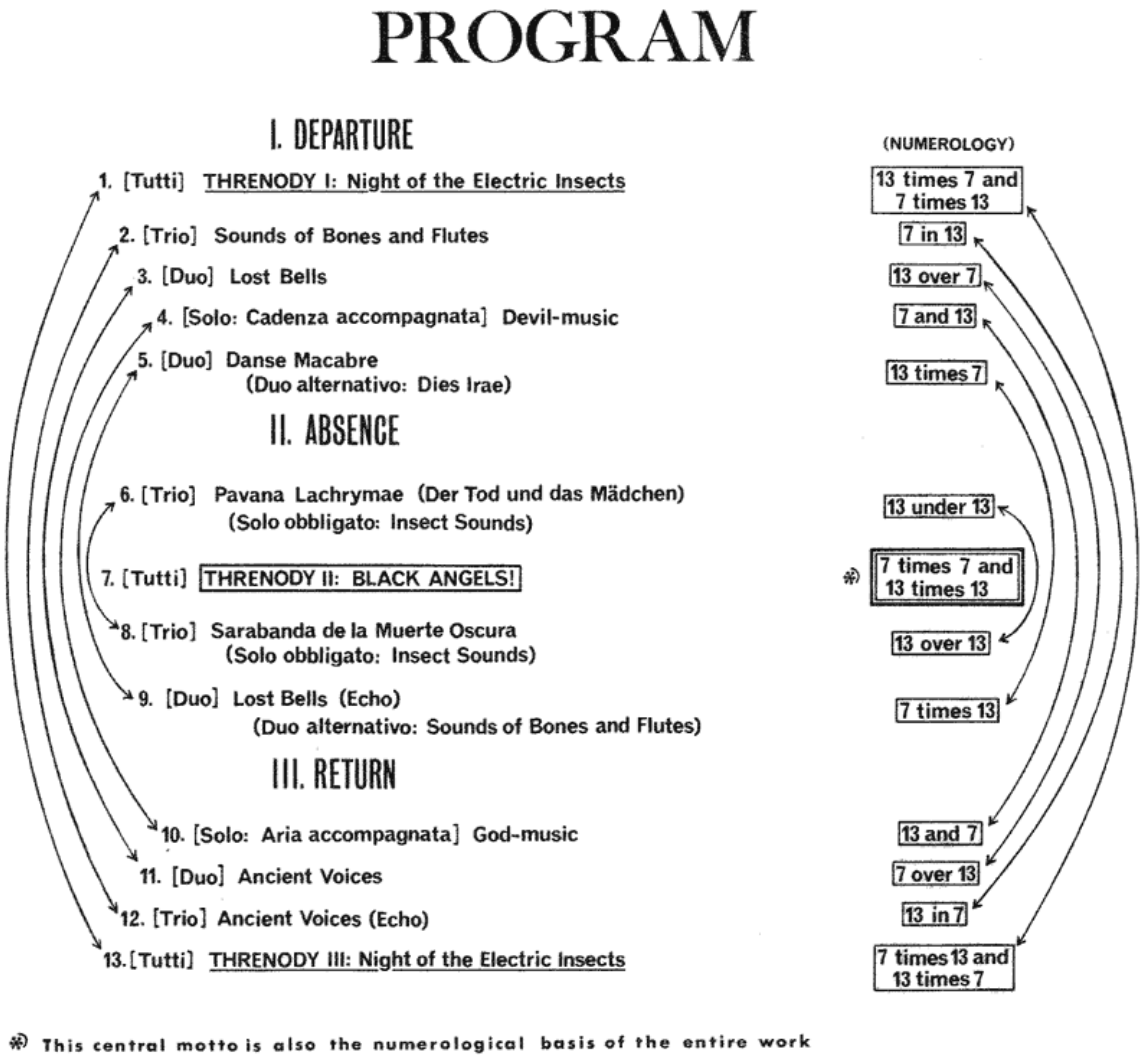

Figure 7. Structure of movements and images of *Black Angels: Thirteen Images from the Dark Land*, electric string quartet.

.

The movement opens with a passage labeled "Pavana" that is a slightly altered version of the "Death and the Maiden" theme that Schubert used in his D minor string quartet, and that Crumb has arranged for three strings instead of four. In the Kronos Quartet recording, an aural space is established that is immediately recognizable as that of a cavernous indoor enclosure—large and empty. Long reverberation times hint at hard, flat

surfaces, perhaps, given the context of a quartet recording, of a church or cathedral. That sense of grand spatial dimensions, however, is disrupted by the loud and strident *tutti* exclamations of the following "Threnody II: Black Angels!" section. At this point, all the action of the piece moves up front—that is, up close to the listener. After the highly active and angsty Threnody, the tranquil, reverberant mood returns in section 8, "Sarabanda de la Muerte Oscura." This section, like the Pavana, is played by a trio within the ensemble, though this time not based on any piece in particular (Crumb calls it "stylistically synthetic").[185] Here, a unique combination of timbres is rapidly juxtaposed: pizzicato strings, voice, harmonics, maracas, and viola *col legno battuto*. Due to their different spectra, these instruments offer competing renderings of the spatial "ground" of Moylan's "host environment." The sense of open space soon returns with the unaccompanied violin's natural and artificial harmonics, as their echoes and reverberations are quite audible. The movement ends with the sound of a bowed tam-tam, a transition to the quartet's third and final movement.

Making an appearance in several images of *Black Angels* is a buzzing, insect-like sound played soft and high in the first violin. In "Absence" this sound occurs in images 4 and 6. These passages are explicitly labeled in the score as "Solo Obbligato: Insect Sounds" and provide a dissonant and unconventional counterpoint to the accompanying trios.

The common practice tonality and the insect obligato found in images 4 and 6 give an ABA' structure to the movement, with the climactic "Threnody: Black Angels!" image in

185 George Crumb. 1972. Liner notes. *George Crumb/Charles Jones, Black Angels for Electric String Quartet/String Quartet No. 6, Sonatina*, with the New York String Quartet, CRI SD 283, 33 ⅓ rpm. Cited in Blair Johnston, "Between Romanticism and Modernism and Postmodernism: George Crumb's Black Angels," *Music Theory Online* 18, no. 2 (June 1, 2012), http://mtosmt.org/issues/mto.12.18.2/mto.12.18.2.johnston.html.

the middle and the final "Lost Bells (Echo)" image functioning as a coda (itself set off from surrounding images by the sound of a bowed cymbal at the beginning and end).

The different images of this movement afford different spatial experiences—and are highlighted by the frequent transitions and juxtapositions among the different timbres. As timbre is a key component of our understanding of space, it is a powerful compositional tool for effecting spatial experiences for the listener. Below I examine some illustrative moments in "Absence" that point out the spatial significance of this spectral dimension of musical sound.

4.2.2 Image 6: Pavana

The movement begins with the image titled "Pavana." The voices in this passage move in a legato fashion from one note to the next, with no sharp attacks. Their soft harmonies offer an expansiveness—an awareness of reverberation highlighted by consonant intervals. The consonant harmonies facilitate the blending of direct and reflected sounds to create the warm sound often associated with string choirs. It is only the first violin's electronic "insect sounds" (with the expression marking "nervous, fleetingly") that alert us to the full dimensions of the space in which the ensemble is located. Not only does the violin have an audible reverberation perceivable as separate from the direct sound, it is also marked apart from the trio by its rapid movement, faint dynamic level, and the timbral qualities indicated by the trill, tremolo and *sul ponticello* indications. In spite of the buzzing insect part, the passage creates a contemplative atmosphere. Certainly, the buzzing alerts us that something is amiss in this particular atmospheric construction, but its soft, meandering line elicits curiosity more than tense anticipation. As in the Miles Davis

examples in Part I, we are offered here a particular mode of listening that comes as a result of the amalgamation of our own listening situation with the spaces afforded by the music and portrayed by the recording: an example of Hansen's "mixed reality." The echoes and reverberations are sounded in our own space as well.

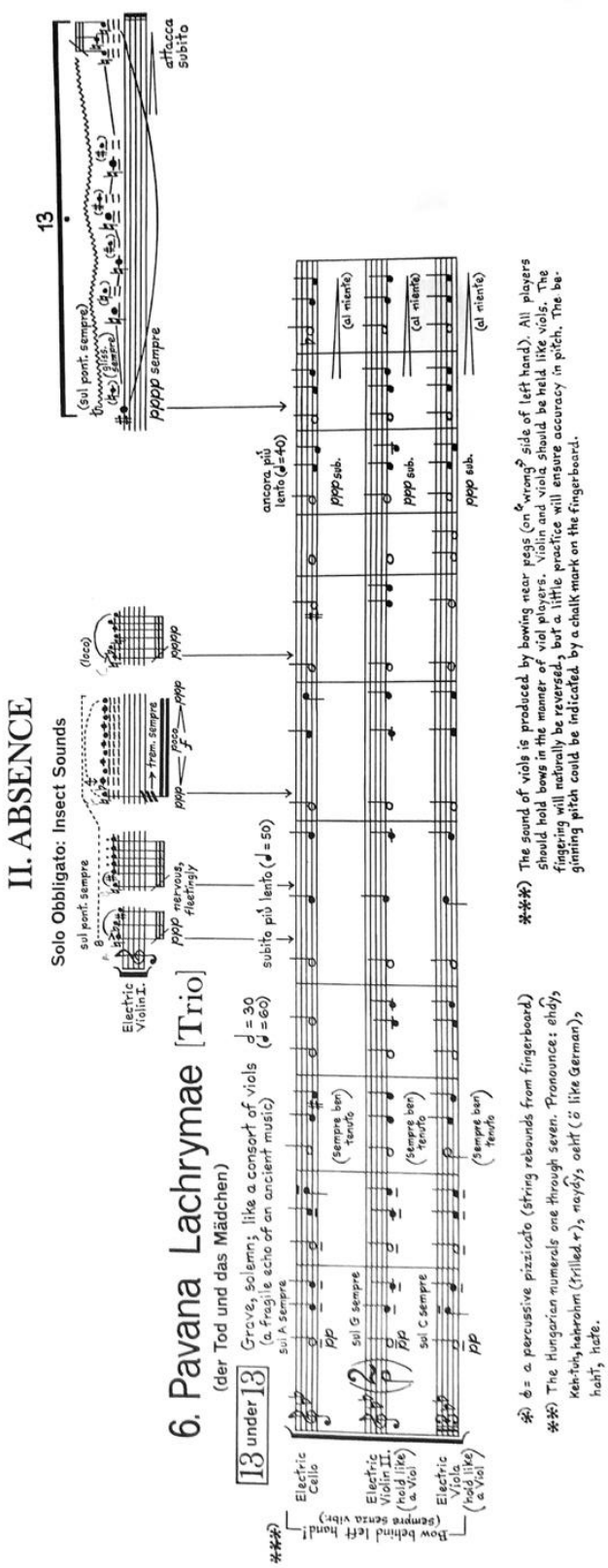

Figure 8. George Crumb, *Black Angels*, Image 6, "Pavana Lachrymae."

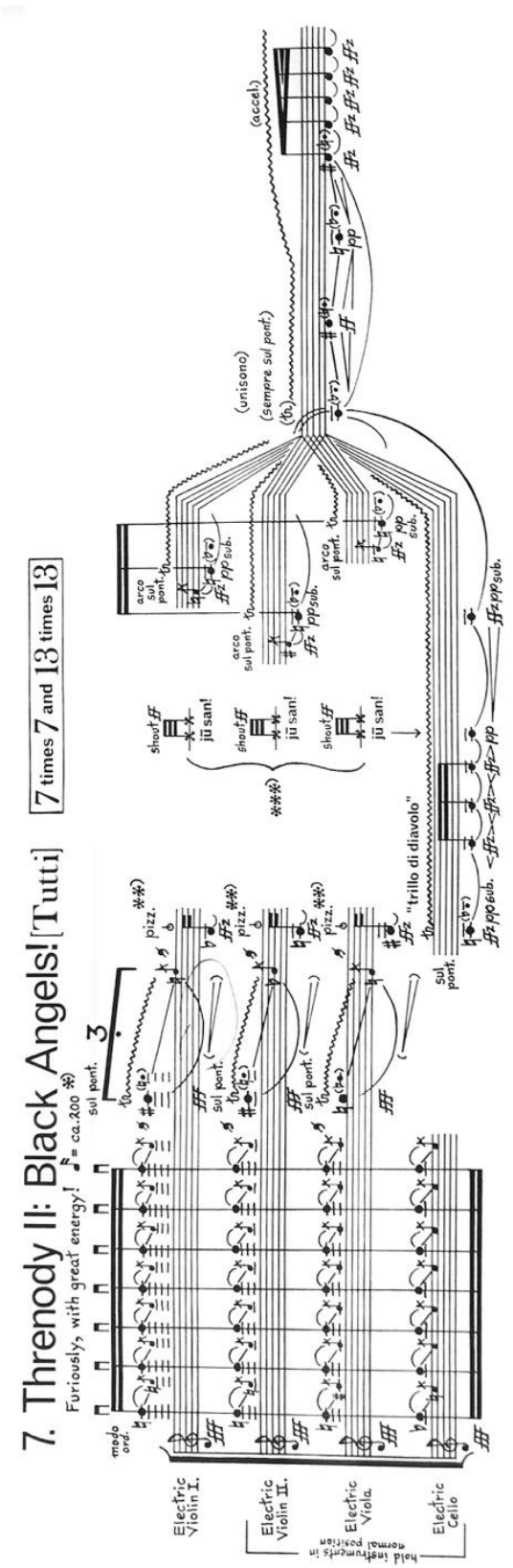

Figure 9. ***Black Angels*****, Image 7, "Threnody II: Black Angels!," system 1.**

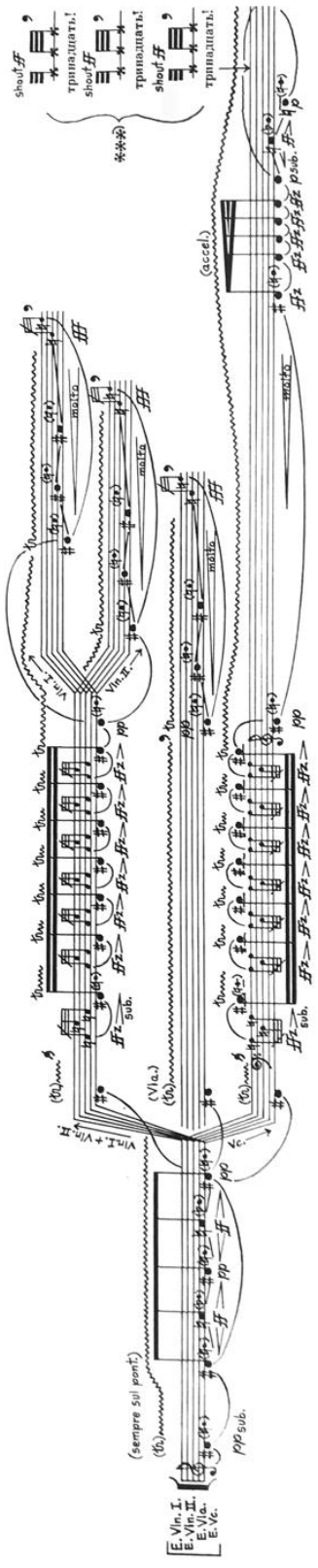

Figure 10. *Black Angels*, Image 7 (cont.), system 2.

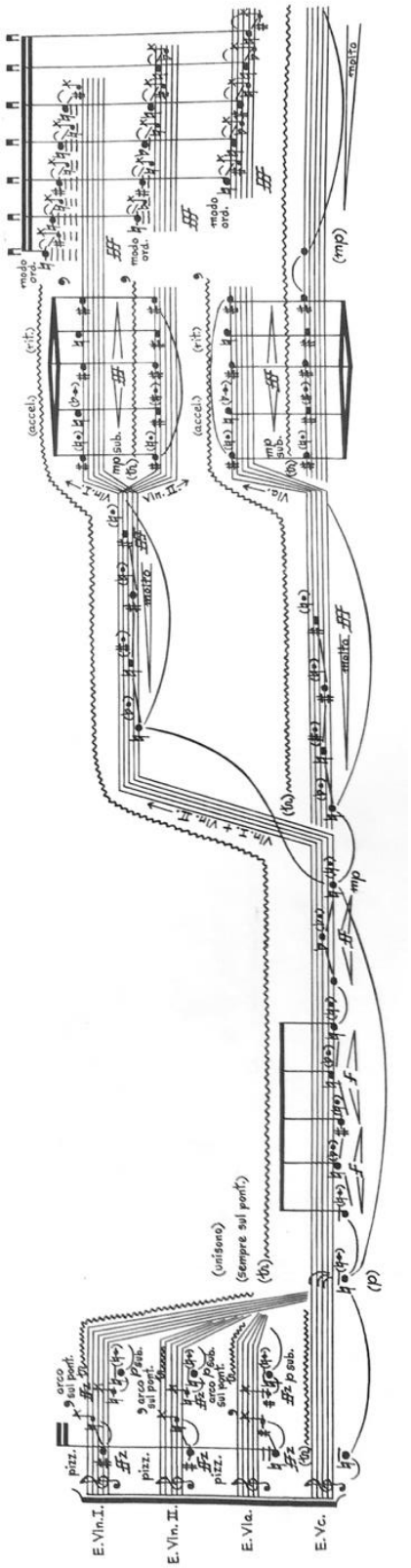

Figure 11. ***Black Angels*****, Image 7 (cont.), system 3.**

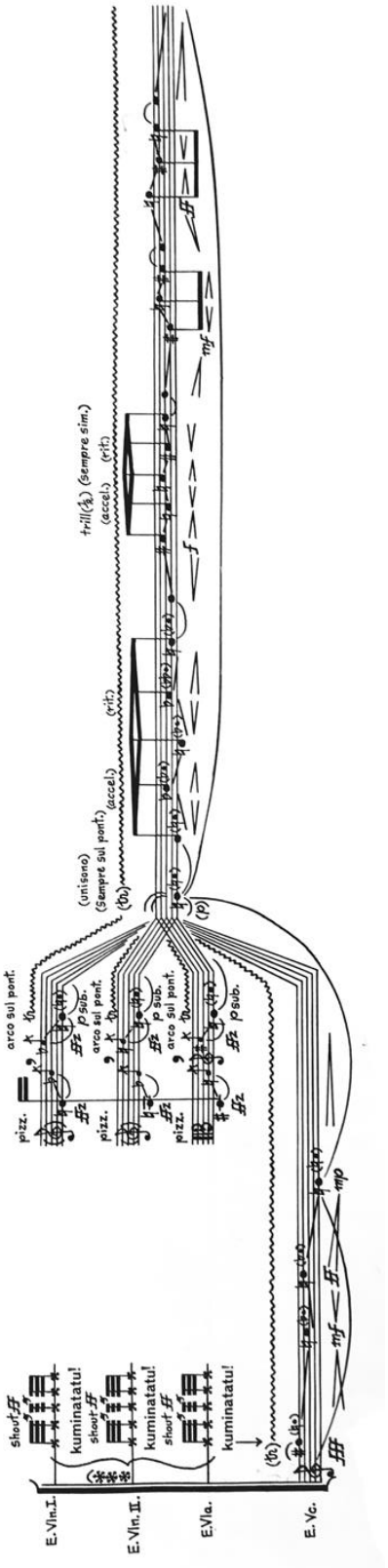

Figure 12. ***Black Angels*****, Image 7 (cont.), system 4.**

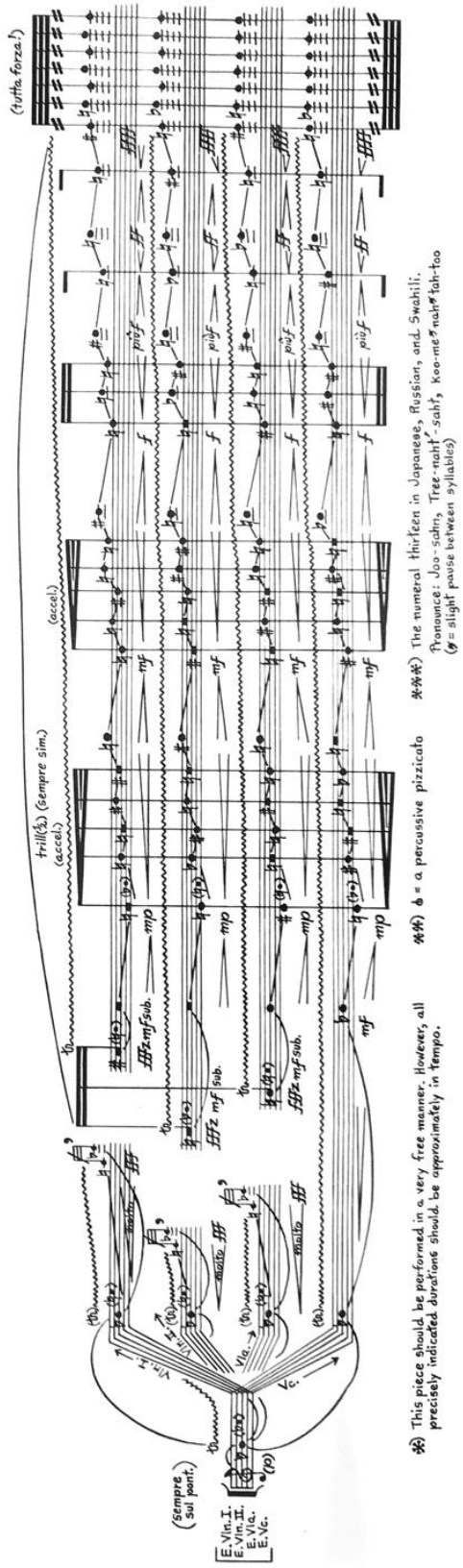

Figure 13. *Black Angels*, Image 7 (cont.), system 5.

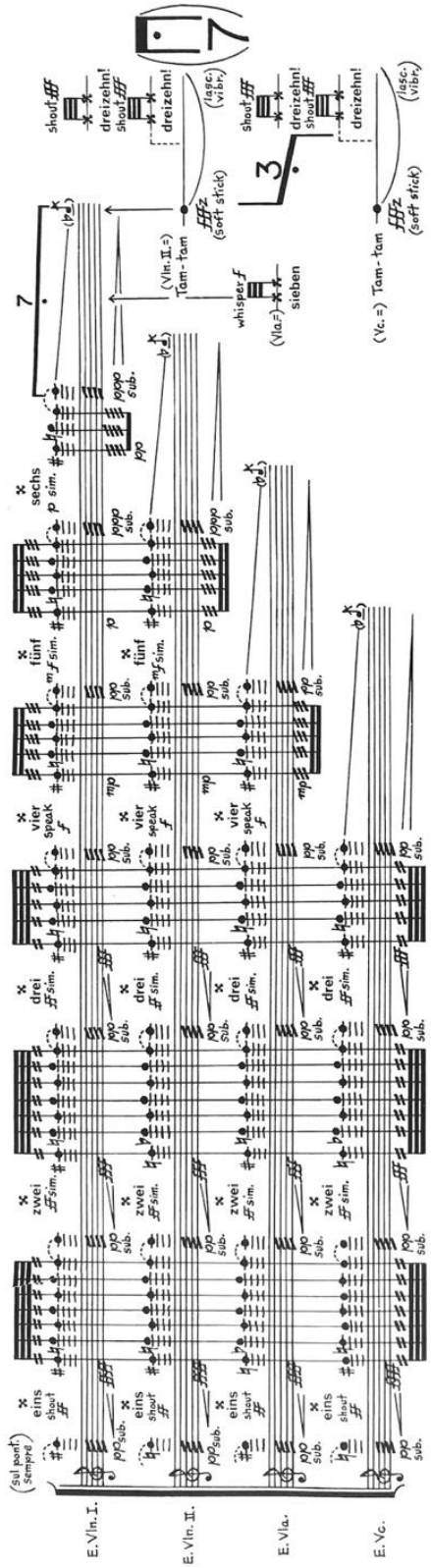

Figure 14. *Black Angels*, Image 7 (cont.), system 6.

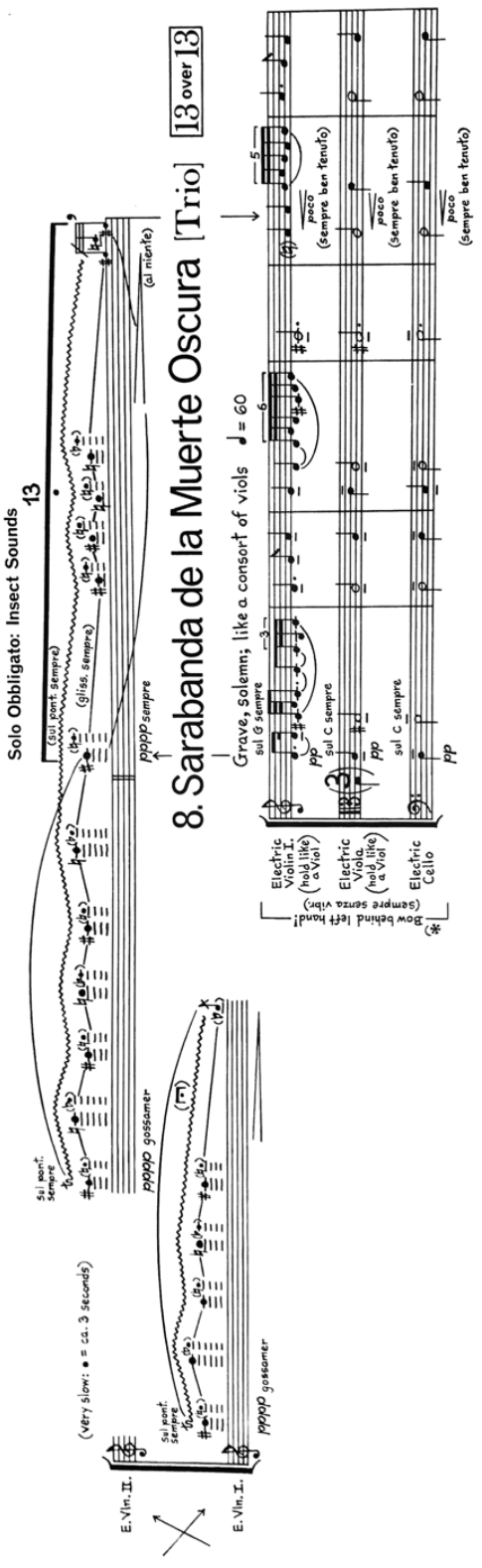

Figure 15. *Black Angels*, Image 7 (cont.), system 7; Image 8, "Sarabanda de la Muerte Oscura."

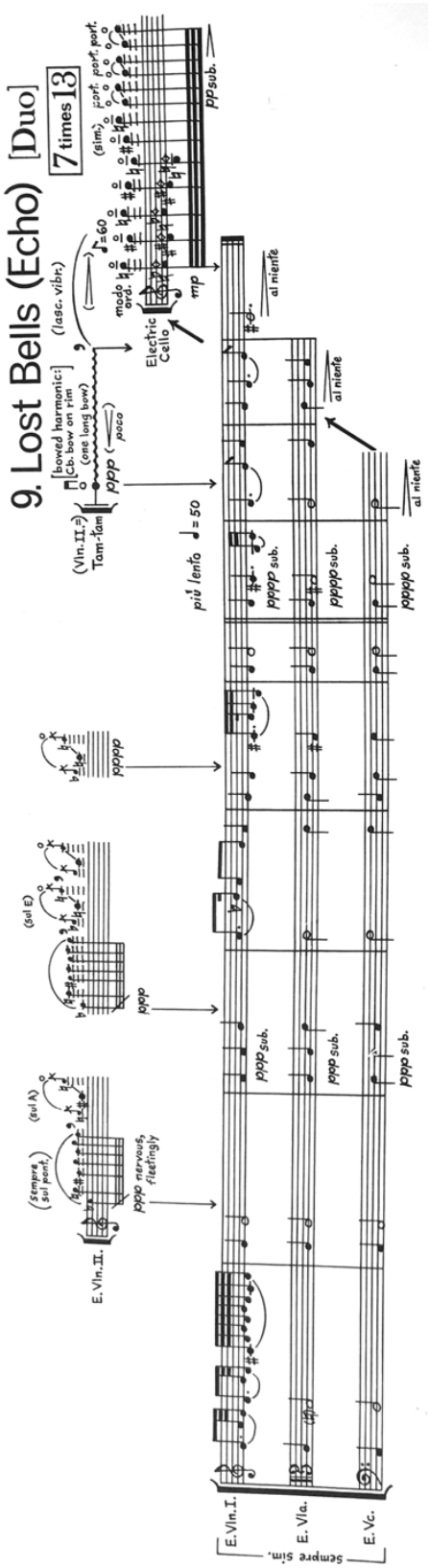

Figure 16. Black Angels, Image 8 (cont.); Image 9, "Lost Bells (Echo)."

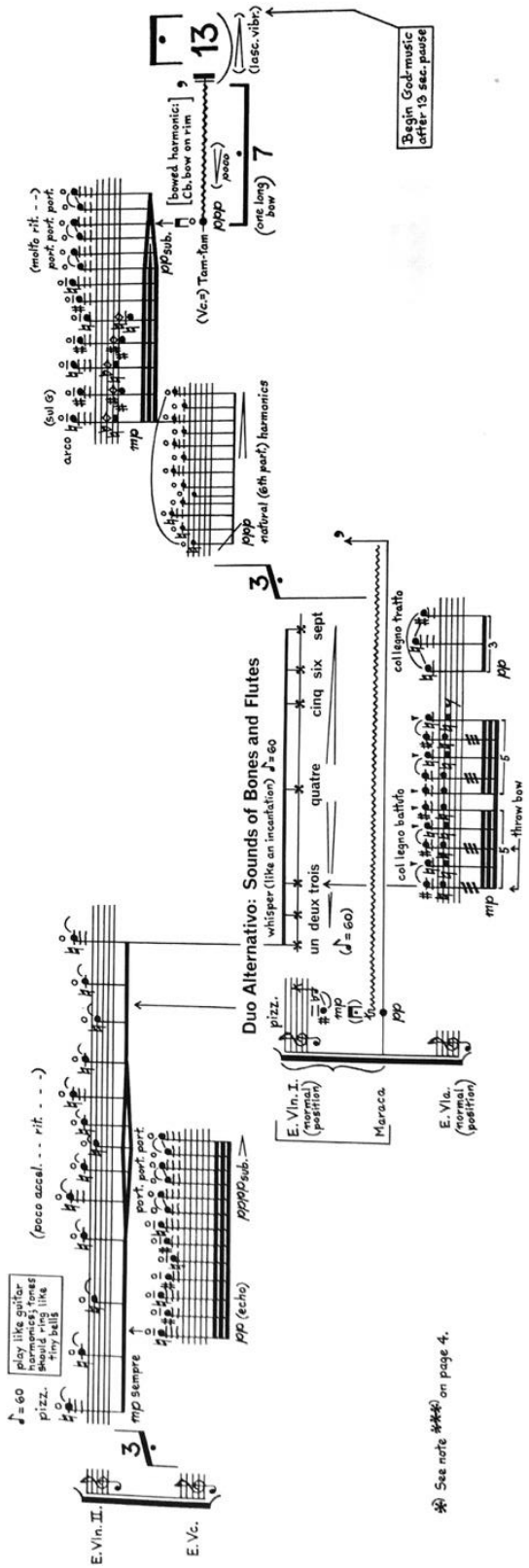

Figure 17. Black Angeles, Image 9 (cont.), system 2.

4.2.3 "Threnody II: Black Angels!"

The arrival of the Threnody makes clear the vulnerability of this mood. Any sense of spaciousness, however, is overridden by the sheer loudness of the beginning of the Threnody. The section begins with seven swift, dissonant, triple-forte pitch glides in each of the strings, outlining tritones beginning on Bb5, E6, F6, B6 (Figure 9). The attacks are strident, with a clearly audible scraping sound of the bow across the strings We are suddenly accosted by which shut down any space for contemplation and force us to attend to a single, repeating sound. The attacks on the strings are so sharp they mask much of the reverberant sound, while their echoes come through in the brief interval between bows. The change is disorienting not simply because it is loud and surprising, but also because of the shift in perspective: from one in which the listener is located in a shared space with the players, to one in which the listener has sounds coming out at them. These glides cannot be heard otherwise than extremely close. The string attacks do not usurp the space evoked by the Pavana, but they do shift it to the background, temporarily out of focus. Such a shift in perspective is not just disorienting, but can register as an aural assault: there is no longer a comfortable distance between sound and listener. What we notice are the echoes of the glides and very little of their subsequent reverberation.

Following the opening seven strokes, there is a three-second glissando in the upper strings ending in a percussive pizzicato. Then the texture thins to one voice and reverberations can be heard as the original "sound object setting" or host environment of the Pavana returns. Of course, that setting never really left, it was simply moved to the background as there were other things to attend to. Such shifting of features to the periphery is an aspect of visual perspective as described by Merleau-Ponty, but it is equally

applicable to auditory perspective.[186] We don't give a second thought to these commonplace moments that break open our spatial frame of reference and enhance our perspective, but they are remarkable nonetheless. When, in a mixed reality situation, the location cues of an alternative spatiality are removed, we are tasked with reassessing the boundaries of competing markers of place, and of the nature of the subject-object relation under these new circumstances. The transition to the Threnody makes clear the vulnerability of our spatial situation as constructed through these mixed reality assemblages.

4.2.4 Tam-tam

At the end of "Threnody II: Black Angels!" (2'55"), after a decrease in dynamics and a thinning of the texture to a single voice (the first violin, with a seven-second decrescendo that starts at *ppp*), two simultaneous tam-tam strikes marked *fffz* obliterate not only the dying glissando of the first violin, but the entire spatial configuration that preceded it (Figure 14). The dimensions of the listening space that had been foregrounded earlier are no longer discernible. The immense sound of the gong surrounds the listener, regardless of the local sound object setting in which the recording is played: the sound is equally in front and behind them and they have no sense of where they are within it. The origin of the sound is obscured by the multiple inharmonic overtones of the instrument. The sound the gong emits is suffused evenly throughout the acoustic space of the recording environment

186 "In vision…I apply my gaze to a fragment of the landscape, which becomes animated and displayed, while the other objects recede into the margins and become dormant, but they do not cease to be there….The horizon, then, is what assures the identity of the object throughout the exploration." Merleau-Ponty, *Phenomenology of Perception*, 2012, 70.

and also, to a large degree, throughout the listener's environment. It is not a matter of being surrounded by the sound, because the sound here has no direction at all: a single source location cannot be identified. During the double tam-tam hit there is a *tutti* unison shout, marked triple-forte, of the word "*dreizehn*." The echoes and long reverberation of the voices indicate a large performance space, yet the sustained ring of the tam-tam does not. The diffusion of inharmonic partials leaves us unable to site its origin in any particular location, or any particular host environment. This is because the ear does not have the phase information it needs to make judgments of location. As a result, the marked echo created by the shout does less to place the shouters in the listening field than it does to further blur their location. The shouters are like one of the processed tracks described by Serge Lacasse in his study of vocal staging: an instrumental character with its own host environment functioning within a larger, staged performance environment.[187] The shout is a sound located somewhere *within* the crash—echoey, but in no place in particular.

A unique feature of the tam-tam that contributes to its immersive quality is the curious relation of its attack to its decay and sustain. The sound of a tam-tam builds after being initially struck, as low frequency energy gets transferred to higher modes of vibration.[188] An initial low roar rapidly expands into a loud, bright shimmer that continues unabated for several seconds before starting a very slow decay. As a metal surface struck with a yarn or felt beater, its attack does not afford the type of early reflections that Moylan identifies as crucial to discerning room size. Recall that Moylan attributes our ability to

187 Serge Lacasse, "Persona, Emotions and Technology: The Phonographic Staging of the Popular Music Voice," *CHARM Symposium 2: Towards a Musicology of Production*, 2005, 11.
188 N. H. Fletcher, "The Nonlinear Physics of Musical Instruments," *Reports on Progress in Physics* 62, no. 5 (January 1999): 760.

sense a host environment to reflected sound, some of which is reflected so rapidly we don't hear it as separate from the direct sound. Therefore, the lack of attack means that yet another cue for placing ourselves within a larger enclosure is unavailable.

Immersion thus manages to frustrate perception of the sonic object. Without the sense of other things around us that provide depth, where we can imagine perspectives other than our own, we lose our purchase on sound-as-object and consequently are left with sound-as-environment.

As a comparison, note the way the cello's harmonics sound at the end of the Lost Bells section, image 9 (Figure 17). Between whispered French numerals and a dissonant sounding combination of artificial harmonics and bowed thirty-second notes are a slurred group of natural harmonics—marked as an appoggiatura—leaping back and forth by perfect fifths among three pitches D5, A5, E6. These notes are set apart from what surrounds them by their smooth and comparatively less abrasive quality. The notes comprise one of the "flutes" of this section's "Sounds of Bones and Flutes." Here the sound is close. The intervallic relationship of a perfect fifth ensures that the pitches have overtones in common. As they lack prominent inharmonic components, overtones simply reinforce each other as they reverberate across the space. The blending together of these reflected tones obscures the distance cues, but in a different way than the tam-tam. The natural harmonics have a preternaturally clear sound, free of noisy, inharmonic high-energy upper partials. All the other sounds in this section have a clear echo and betray the dimensions of the space—not only the pizzicato notes, but also the stage-whispered numbers, the viola's bouncing and scratching *legno battuto*, and the odd combination of natural and artificial harmonics just before the bowed tam-tam at the finish of the

movement. As a result, we sense a closeness, a distance, and then an effacement of those measures. While there are always some spatial markers inscribed in recordings, they do not carry sole responsibility for music's spatial character. A voluminous, empty, vaulted church can play host to an array of spatial experiences, not just one.

4.2.5 Insect Sounds

Further insight into how timbres can color our spatial perception is offered by what Crumb labels "insect sounds"—obbligato passages played by the electric violin I in the Pavana of image 6 and passed over to the electric violin II part in the Sarabanda of image 8. The entrance of this insect sound in the fifth measure of image 6 is the first challenge to the movement's initial spatial frame (Figure 8). The sound makes its appearance in the same space as the trio, but farther away, hovering, not blended.

The first violin imitates an object in motion: its glissandi replicate the Doppler effect generated by an insect's meandering flight path relative to a stationary listener. But this is not the only aspect of the violin line that separates it from the other voices. Blair Johnston details how the pitch class content of the insect sounds is complementary to those used in the pair of older style trios that appear in this movement—a compositional technique that serves to further detach the obbligato stream from the other voices.[189]

True to Johnston's description of these images' triadic, homophonic framing apparatus, the obbligato lines hover above the historical styles rendered below (Johnston calls them "pastiches"). In the case of the Pavana, in the first four measures, before the

189 Johnston, "Between Romanticism and Modernism and Postmodernism."

insect's entrance, the trio carries meaning through stylistic allusion, via its melodic, harmonic, and, to a certain degree, instrumental similarity to its early 18th century model. As a result, we might not find ourselves paying too much attention to structural aspects of voice leading and harmonic progression and instead simply drift along with the music. But the obbligato frustrates this mode of listening until it ultimately fades away at the final instant of the section. As it enters, the obbligato also effects a shift of attention away from the top voice in the trio and toward itself. Regardless of its stated role in the score, the obbligato's appearance, though wholly out of character with respect to the trio, and in a very high register, pushes the trio down into an accompanying role. Here again, the spaces outlined by the music affect the way we direct our attention. Like the opening strokes of the Threnody, the solo obbligato has the capacity to negate the dimensions of the virtual setting, demonstrating that more than just volume is implicated in the process of refiguring our situational environment. Our spatial awareness is heightened in these moments which disrupt Hansen's "mixed" variety of the virtual. Amplitude is just one way among many to alter spatial parameters and afford a mixture of spatial experiences. The attentional stance the music affords also plays a direct role.

What then is the primary object of musical interest here in these two trio sections? To put the question another way, what is foreground and what is background? Or, with respect to sound studies, what is the "sonic object setting" and what is the sonic (or musical) object? In one way of hearing this passage, the slow, chorale-like progression of the trio is background to a soft yet conspicuous, higher pitched, highly active figure in motion. Mimicking a familiar homophonic texture, we are apt to comprehend the first violin part as a melody line, as an object of focus and attention, that is, as foreground. Yet

we also notice this violin is completely independent, unfettered and unrelated to the steady movement of the other voices in the trio. Is it an intrusion to be ignored, like extraneous noise from audience members at a performance? The breaks in the first violin's line allow us to concentrate more carefully on the Pavana—and we note it has its own melodic line in the top voice and is itself homophonic. While we can't necessarily locate the source of the cello-insect, its incongruous juxtaposition forces us to take a broad perspective in order to have the maximal "view" in this tableau. The tension between these two competing perspectives is responsible for the passage's spatial expression, giving a sense of depth to what is otherwise merely a simulation of the type of "churchy" setting identified by Zayaruznaya.

4.2.6 Speech

There is one timbre in this movement that is peculiarly striking in terms of effecting a sense of proximity: the voice. Crumb calls upon instrumentalists to variously speak, shout, and whisper throughout the piece. Not sung and never sustained, the words function as percussion, much like the s*forzando* pizzicati of Threnody II. The words used in the piece are numerals, recited in different languages: Japanese, Russian, French, German, and Swahili. Interestingly, the attention we direct toward these utterances is not wholly determined by their timbre, dynamics, or the musical contexts in which they occur. As discussed above, spoken words that we recognize galvanize our attention; There is a strong tendency to hear them unthinkingly *as* words, that is, as ordinary language. In contrast,

words we do not understand, or words that are sung, shouted, or whispered, we attend to differently. We understand them as much by their timbre as their lexical meaning.

In a work so dependent upon its timbral construction as *Black Angels*, we are more likely to interpret such utterances as musical elements. Words that are spoken, however, stand out. Especially when spoken softly by a single voice into a nearby microphone: less energy in the direct signal means less energy in the reverberations, and in a resulting early sound field that consists primarily of direct sound, indicating a very close source near the listener's body. Such words can disrupt our mode of attending to the music and call us into a personal relationship with the performer. Stephen Feld describes just such an experience while listening to a story recited by one of his Kaluli interlocutors. At one point the speaker breaks the narrative to point at Feld's microphone and compare it to the snout of his canine protagonist. Feld points out the striking sense of intimacy that arises from such direct address: "What new space of intimacy, what heightened juxtaposition of bodies, of sensualities and desires has Ganigi [the storyteller] now opened up?"[190] Feld's sense of the listening space has changed through this act. Similarly in *Black Angels*, Crumb's use of speech can also refigure the space of the listener. An intimacy can be created not just through the use of microphone technique, but through the composer's use of dynamics and timbre and language.

Starting at the climax of Threnody II, (Figure 14; 2:36) a string of numbers is recited in between ferocious trills. Very little familiarity with German is required to recognize the series "*eins*, *zwei*, *drei*, …" that the players shout out between trills. When the music arrives at the word *sechs*, however, the music seems to come to a halt. This word is neither shouted

190 Steven Feld, "They Repeatedly Lick Their Own Things," *Critical Inquiry* 24, no. 2 (1998): 456.

nor whispered, and it is uttered by only one performer. Though a gradual decrescendo is accomplished through both diminishing dynamics and a thinning of voices, the text performed here on the Kronos Quartet recording stands out as an utterance outside the dynamic gesture. Then moving on to the word *sieben*, which is whispered, the text moves back into something more musical, and enchanted. And enchantment is indeed the effect specified by Crumb when he later uses whispered speech in the duo alternativo of image 9: "Sounds of Bones and Flutes." Above the words he writes the instruction, "whisper (like an incantation)."

While the words "jū san" and "kumi na tatu" are likely to be perceived by the Western listener according timbral and rhythmic affordances (and probably the word "тринадцать," as well), this is not the case for the words familiar to the listener—especially when they are not shouted but spoken softly, by a single speaker, unencumbered by the distancing effects of echo. The music does not merely symbolically represent or even index intimacy. It is a direct address that draws the listener into a particular situational context proposed by the speaker. It is not only the contrast in volume, and accompanying reverberation and echoes, that make "*dreizehn*" (the final shouted "thirteen" at the end of the "Threnody II" section) affectively different. It is also the fact that the listener's situation is different: they are not an interlocutor. The listener is rendered passive, subjected to the shouts rather than in dialog with them. What is remarkable about "*sechs*" is how the listener's stance toward the music is transformed in an instant. In the Thenody II section, it seemed as though the cavernous host environment had shifted to the background with the arrival of the seven *tutti* down-bows. Here, that environment seems to have disappeared altogether, because the music draws upon a different contextual frame. More than simply

sounding close, the entire music listening context is refigured by drawing upon a listener's capacity for sensing intimacy.

4.2.7 Maracas

Another moment that fosters a closer, more proximal, and also more intimate relation to the music is comes at the end of the movement, but that closeness is accomplished by a different means than vocal effects or recording techniques. In image 9, "Lost Bells (Echo)," a unique combination of timbres is rapidly juxtaposed: pizzicato, voice, harmonics, maracas, and viola *col legno battuto* (Figure 17; 4:46). The rolled maracas blend with the final pizzicato notes of the violins, significantly altering the spatial expanse indexed by their reverberant harmonies. When the maracas begin, they cover up the reverberations of the remaining pizzicati, with the effect of "flattening" the music, removing any cues that point to the surrounding acoustic performing space. The sense of open space soon returns with the unaccompanied violin harmonics. These particular acoustic effects can be specified in the score; however, they are not always predictable, as Rebecca Leydon demonstrates with her comparison of two recordings of *Black Angels*.[191]

Leydon's study is also focused on Crumb's use of timbre. She demonstrates how different sounds can become fused into a single timbre through an interaction of the instructions provided by a score and the choices made by the performers. As an example, she notes the return of the Sarabanda in image 13. Leydon says that listeners make choices as much as performers do, and we can certainly imagine that the tiny head movements

191 Rebecca Leydon, "Clean as a Whistle: Timbral Trajectories and the Modern Musical Sublime," *Music Theory Online* 18, no. 2 (June 2012), https://mtosmt.org/issues/mto.12.18.2/mto.12.18.2.leydon.html.

detailed by Xie can help determine whether such a fusion happens for listeners or not. Leydon, too, understands this timbral perception as a "binding" together by an engaged listener of separate strands of the music. Her analysis suggests that this grasping together is also the source of the work's affective power. Leydon writes that as violin harmonics split off from the textural weave to form their own layer during the second movement, they set up a contrast between the ethereal and the material—"timbral evidence of bodies and non-bodies"—and thus, she claims, a contrast between worldly existence and a "super-corporeality…. a physical body that has become sharply separated from its own spiritual essence." Later, in the third movement, a curious sounding return of the Sarabanda melody is tapped out on the strings by thimble-covered fingertips. Whether the timbres blend or are kept separate can influence the meaning that emerges for the listener: in Leydon's interpretation, whether matter and spirit can be reconciled or if they are fated to remain separate.

Leydon analyzes two different recordings to illustrate how the same notated figure can result in distinct timbral effects. In one of them (the Kronos Quartet recording I have been analyzing here) the melody of the Sarabanda is all but inaudible. But in a recording of the same passage by the Carnegie-Mellon Contemporary Ensemble, Leydon notes that the Sarabanda "sings through" the gauzy trickling of the thimbles. A blended hearing is offered, yet it is not one the listener is obliged to take up.

> In fact, Crumb's performance directions can just as easily create a fused chimeric timbre here, or, crucially, one that we may choose to hear as blended, as the Sarabanda tune *singing through* the (paradoxically) "noisy" harmonics of the thimbles. The cello's ghostly interjections encourage this orientation as they anticipate and echo fragments of the Sarabanda melody. Alternatively, we may choose to separate the strands into two discrete layers: plinking harmonics and a faint melody.

Similarly, with respect to image 9, a particular performance configuration or recording situation might result in the maracas and strings being kept distinct. This, of course would articulate the listening space in a different manner. Yet regardless of the characteristics of a piece a performer chooses to emphasize, the listener will always have an active role in what is ultimately perceived.

Leydon's analysis points us toward another spatial concept to consider: atmosphere. Atmosphere has been theorized as an aesthetic phenomenon by Gernot Böhme, and his description explicitly refers to the twinned determinants of timbral perception mentioned by Leydon: performers and listeners.

> Atmospheres are something between subject and object: one might call them quasi-objective feelings that are indeterminately diffused in space….. However, insofar as they are nothing without a perceiving subject, they also have to be called subjective. Their value lies, precisely, in this in-between state, bringing together what was traditionally separated into the aesthetics of production and the aesthetics of reception.[192]

What if this quasi-objective nature of atmosphere could be rendered more concrete? What would that sound like and how would it be accomplished? The music of Éliane Radigue, which I will examine in the following chapter, provides an example of one way such the mechanics of atmosphere might be revealed musically.

192 Gernot Böhme, *Atmospheric Architectures: The Aesthetics of Felt Spaces*, trans. Tina Engels-Schwarzpaul (New York, NY: Bloomsbury Academic, 2017), 125.

4.3 Conclusions

All the analyses thus far have focused on moments that depend upon the interaction of music and listener and space for their realization. As Marion A. Guck puts it, "the point is not identifying configurations of notes but showing how my experiences are elicited by the ways in which the configurations come together for me and change me as I respond to them."[193] In the Miles Davis track "Rated X," the abrupt change of spatial volume occupied by the organ can only occur when not only the rhythm section is cut out, but when all the attendant studio reverberations are silenced as well. Like gas in a vacuum, the organ sound suddenly expands to occupy the listening space abruptly abandoned by the other instruments. In the Clerks' Group recording of Ockeghem's mass, resonant properties of the recording space sounded out by the voices magnify the volume of space occupied by the music through changing the focus of attention from pitches and intervals of a contrapuntal passage to the acoustic blend of the harmonic frequencies of a sustained triad. In the Kronos Quartet recording of Crumb's *Black Angels*, I highlighted the disorienting effects of the tam-tam, an instrument that produces inharmonic partials, thus lacking the sonic cues we use to sense the dimensions of the space around us. Thus it frustrates our ability to situate ourselves with the spatial dimensions projected by the recording and as I discussed in my analysis of Davis's "Maiysha."

Through analyzing the timbres that Crumb uses in *Black Angles*, I have demonstrated the ways in which the recording environment can have an impact on how we interpret the musical sounds we hear. Reverberation that originates in the recording's host

193 Marion A. Guck, "Who Counts?" (paper presented at the joint meeting of the American Musicological Society, Society for Ethnomusicology, and Society for Music Theory, New Orleans, LA, November 3, 2012). Cited in Eidsheim, *Sensing Sound*, 190.

environment affects how sounds blend together for listeners. But other musical sounds, like the maracas and tam-tam, thwart some of the common effects of reverberation, and complicate both the identification of the dimensions of the host environment as well as our ability to locate ourselves spatially within our own listening environment. The tam-tam's inharmonic vibrations do not allow our ears to detect phase differences, or distinguish between direct and reflected sound. Thus, a certain type of relationality to our environment is denied us and are sonically cast adrift. Conversely, adding maracas to the texture of "Lost Bells (Echo)" masks the reverberant sound made by the notes played *pizzicato* and the viola's passage *col legno battuto*. The voices, too, have their reverberations obscured by the soft swirling of the maracas' contents within their shells. The acoustic characteristics, therefore, of neither the recorded performing space nor the ultimate listening space are fixed by those spaces' physical dimensions. Not every sound recorded in a large hall will necessarily be heard as if played in a large hall. Because of this fact, we can safely assert that music has the capacity to alter our sense of the size of the spaces we inhabit, and that through the use of certain combinations of instrumental timbres, the distance from the listener of the sources of musical sounds can be manipulated as well.

Chapter 5: Éliane Radigue, "Kailasha"

5.1 Space and Texture

Take a ping-pong ball. Cut it in half. Place the halves over your eyes and *voilà*, you are looking at a Ganzfeld. The Ganzfeld is a key concept in the work of German Gestalt Psychologist Wolfgang Metzger and was discussed by James J. Gibson in *The Ecological Approach to Visual Perception*.[194] Metzger first devised this phenomenon by adjusting the lighting that fell on a white wall. By lowering the light, he could make the wall fade from view so that all observers could see was an indistinct glow. Metzger claimed that what subjects saw in such cases was "space": the third dimension that was left over after the two-dimensional wall had disappeared. Gibson, however, claimed that what they saw was *nothing*.

Gibson described some experiments of his own with the Ganzfeld. He constructed a set of goggles for his subjects to wear that would produce a completely homogenous field of illumination (using the aforementioned ping-pong balls). The subjects could see light, but nothing else: no shadows, no gradations of brightness—just luminosity. The only apparition available to the viewer is a completely featureless and, indeed, invisible, landscape. The world that appears to the viewer under such tightly controlled conditions is a featureless landscape. Gibson claimed that this uniform distribution of light offered a glimpse of what he considered one of the essential components of the environment: the "medium." According to Gibson, the medium is ordinarily inaccessible to perception. Its

[194] Wolfgang Metzger, "Optische Untersuchungen am Ganzfeld. II. Mitteilung: Zur Phänomenologie des homogenen Ganzfelds [Optical Investigations of the Ganzfeld. II. Toward the Phenomenology of the Homogeneous Ganzfeld]," *Psychologische Forschung* 13, no. 1 (January 1, 1930): 6–29, https://doi.org/10.1007/BF00406757. Gibson, *The Ecological Approach to Visual Perception*.

role in his ecological model is to serve as the ether that allows the visual signals emanating from material objects in our surroundings to reach our organs of perception. Gibson inferred that, given its featurelessness, the Ganzfeld could neither be determined as two-dimensional nor three-dimensional. In fact, he claimed, one could not really say whether it was dimensional at all.

Gibson maintains that what prevents the appearance of dimensionality under such conditions is the lack of texture. Texture, he claims, appears to a seeing organism as patterns of variation in the intensity of ambient light that reaches the eye—variations in the otherwise constant and consistent reverberation of light in all directions as it bounces back and forth among surfaces in the landscape. This patterning he terms "the ambient optic array" and it is what makes the visual perception of surfaces possible. Furthermore, he says that what the eyes are able to detect are not objects, but rather illuminated surfaces. Thus, the existence of textured surfaces and their varying levels of reflection are necessary for visual perception. There is yet another necessary condition for vision: movement. "A surface," Gibson writes, "is seen when the array has structure, that is, differences in different directions."[195] Only an organism capable of locomotion, therefore, is capable of vision. Without the capacity for assuming different perspectives, seeing cannot take place.

When faced with a Ganzfeld, there are no differences in different directions, and therefore surfaces cannot be seen. Hence Gibson's verdict that Metzger's subjects saw nothing, and further, that "depth was not present in the experience but missing from it."[196]

195 Gibson, *The Ecological Approach to Visual Perception*, 151.
196 Gibson, 151.

Textured surfaces, then, are not only what an organism is able to see; they are also what furnish an organism's environment with depth.

There is a sonic analog here, a contrast between a featureless field of sound and one that is somehow textured, a space in which an organism's movement would reveal differences in the perceived sounds. The creation of depth through texture is central to Gibson's approach to vision, and is also relevant to a discussion of depth as it is perceived through listening. In listening, too, different patterns that emerge through movement also provide space with depth.

Depth is a key concept in phenomenological investigations of embodied perception. Scott Marratto describes it as the fact that the objects we perceive appear to us as either near or far, rather than set alongside one another in a two-dimensional tableau. As a result, things in our environment "appear as somehow enveloped within one world," an appearance that gives rise to a sense of voluminosity.[197] The significance of this definition of depth is that it entails the perspective of a body with motor abilities for which things can appear as being near or far. Because depth is revealed through motion, it is a phenomenon of change and difference not bound to binocular vision. Similarly, audible depth does not rely solely on distant echoes, relative volume, or ITDs and ILDs.[198] We have noticed this earlier in the exploration of works by Ockeghem and Crumb. In those pieces, I noted that the music produced spatial effects that did not rely on stereophonic sound or on a specific arrangement of musicians. The same is true for the electroacoustic music of Éliane Radigue: the pulsing eddies of sound that are characteristic of her work emerge from the physical

197 Marratto, *The Intercorporeal Self*, 28.

198 That is, Interaural Time Differences and Interaural Level Differences, consequences of binaural hearing discussed above in the previous chapter.

engagement of the listener within the space in which it is heard, and is not the result of stereo imaging techniques. The full force of her music can only be appreciated through the embodied response of a listener in motion.

Our ears can locate sounds in relation to one another through a variety of mechanisms. Some of these are astonishingly subtle, such as the ability to detect phase differences between signals reaching the left and right ears. But there are also sounds that can frustrate the ears' efforts to discern their point of origin. They may seem to surround a listener and suffuse her environment so completely as to render other perspectives unattainable, and other sounds dislocated. And when such sounds without origins also seem to lack texture, that is, when movements of head or body fail to reveal differences in sound, we have a situation where we may well feel ourselves immersed in that sound. We can think of this immersion as a sonic Ganzfeld. We detect sound, but it is so uniformly dispersed through the medium that it lacks direction and takes on an all-encompassing quality, as in Metzger's experiment with the dimly illuminated wall. But we would not say that we heard nothing, as Gibson claims Metzger's subjects did. We still hear the sounds that we are immersed in as having their own character. For example, unlike the illuminated fog that Gibson equates with a Ganzfeld, we can hear a tam-tam as a tam-tam. Of course, our ability to identify the sound does not mean that a tam-tam crash is not a Ganzfeld. It is not hard to imagine, say, a Ganzfeld of blue light—a Ganzfeld that we perceive as something other than simply nothing.[199] But the sound of the tam-tam has a quality the visual Ganzfeld

[199] David Morris in *The Sense of Space* describes an art installation by James Turrell in which a small opening, like a window, is cut into a wall that looks into an empty room uniformly lit with ultraviolet light. He states that what an observer sees is "a strange blue light." Morris concludes: "Texture is lacking in Turrell's rectangle of light, which is a no-place, and does not move when we move around it."[199] Morris, as well,

does not: it changes through time. The sound evolves as it perdures and one may, for instance, occasionally notice higher partials that become more salient than lower ones, or vice versa. The attack, sustain, and decay phases of its sonic envelope are neither consistent nor linear. But, as noted in the discussion of *Black Angels*, the tam-tam's chaotic modes of vibration inhibit our ability to hear other sonic objects that may be present in depth, that is, in a spatial relationship to one another. Radigue's music provides a subtle yet distinct contrast to these immersive moments in Crumb. Instead of a dense suffusion of sound, Radigue's music has a wavering presence that provides the space itself with texture.

5.2 Analysis: "Kailasha"[200]

5.2.1 Introduction

> What we're trying to do is to surround the audience with a sound, so that when you turn your head you hear something different. The sound is everywhere. There is no stereo; the stereo is everywhere.
>
> — Éliane Radigue[201]

The performance aesthetic at work in Radigue's music is radically opposed to one that values uniformity of sound through a performance space, and that, moreover, enforces that uniformity through conventions such as seating policies, where listeners are required

subscribes to the Gibsonian view that textureless entities are invisible, yet he allows for the color of the light to be perceived.[199] Morris, *The Sense of Space*, 2004, 180.

[200] Éliane Radigue, "Kailasha," Chapter 2 of Trilogie de la mort, Experimental Intermedia Foundation XI 119, 1994, compact disc. Available on Spotify: https://open.spotify.com/track/2aSwsrvPkO86rrYIPSkLyO; https://www.youtube.com/watch?v=lkL-84yRML8.

[201] Paul Schütze, "Surround Sound," interview with Éliane Radigue, October 1, 2011. https://frieze.com/article/surround-sound.

to sit in chairs facing the stage. Radigue crafts variability into her music, ensuring that it not only is dispersed everywhere, but that it is *different* everywhere—a result of her choice of instrument and spare style of composition. Her use of the term "surround" to describe this variable diffusion contrasts with how it is used as a technical specification in audio engineering. Rather than a two-dimensional arrangement of sounds along a horizontal line, Radigue refers to sonic characteristics that can reveal the sense of depth that underlies our spatial experience. Essentially, she is saying that she does not want her sounds to give the impression that they emanate from a phantom source. Instead, she wants them to exist as non-localizable sounds that have the ability to envelop the listener. In addition, the quote identifies a further distinction that the word "everywhere" tends to obscure: the sounds are not "everywhere" as a pervasive, uniform substance. Instead, as Radigue notes, "when you turn your head you hear something different." Her language here echoes Gibson's definition of texture, suggesting that what her music contributes to the shaping of space can be considered in similar terms.

Radigue's music is distinguished from "spatial music" works such as Stockhausen's *Gesang der Jünglinge*, discussed in chapter 1, that are concerned primarily with specifying the direction from which sounds come to the listener. In the literature on popular music and sound engineering, this location information is a component of the "staging" of an individual sound source. Staging is a technical term that refers to the studio practice of mixing vocal and instrumental sounds in stereophonic recording (as well as for 5.1 and other surround-sound formats) so that they seem to originate from a specific site relative

to the listener.[202] Such locations are specifiable along a lateral axis and, in surround sound systems, a sagittal axis, with the addition in some format of a vertical axis as well. The spatiality of Radigue's works, however, is not dependent upon such technology. Instead, spatiality arises out of the enhanced capacity of electronically generated sounds to exhibit interference patterns that undermine our ability to identify the location of their sources. The music's component sounds are all produced by a single electronic instrument, and are fixed, performed, and heard via electronic recording technology. As such, Radigue's music exhibits features that are typical of other recorded works in general and of electronic music in particular, specifically, the reliance on timbre and texture to create musical structure. [203]

[202] For Serge Lacasse, staging includes other aspects of the sound as well, but is still based primarily upon the use of reverberant characteristics as markers of location and performance environment. Serge Lacasse, "Voice and Sound Processing: Examples of Mise En Scène of Voice in Recorded Rock Music," *Popular Musicology Online*, no. 5 (2000), http://www.popular-musicology-online.com/issues/05/lacasse.html. See also Moylan, *Understanding and Crafting the Mix*, 2007, and Ruth Dockwray and Allan F. Moore, "Configuring the Sound-Box 1965–1972," *Popular Music* 29, no. 2 (2010): 181–97. Stockhausen uses the term "panorama" to describe the lateral range of the stereo field and the arrangement of instrumentalists within it. Karlheinz Stockhausen and Jerome Kohl, "Electroacoustic Performance Practice," *Perspectives of New Music* 34, no. 1 (1996): 74–105; 6, 8.

[203] William Moylan, "An Analytical System for Electronic Music" (D.A., Ball State University, 1983), 9.

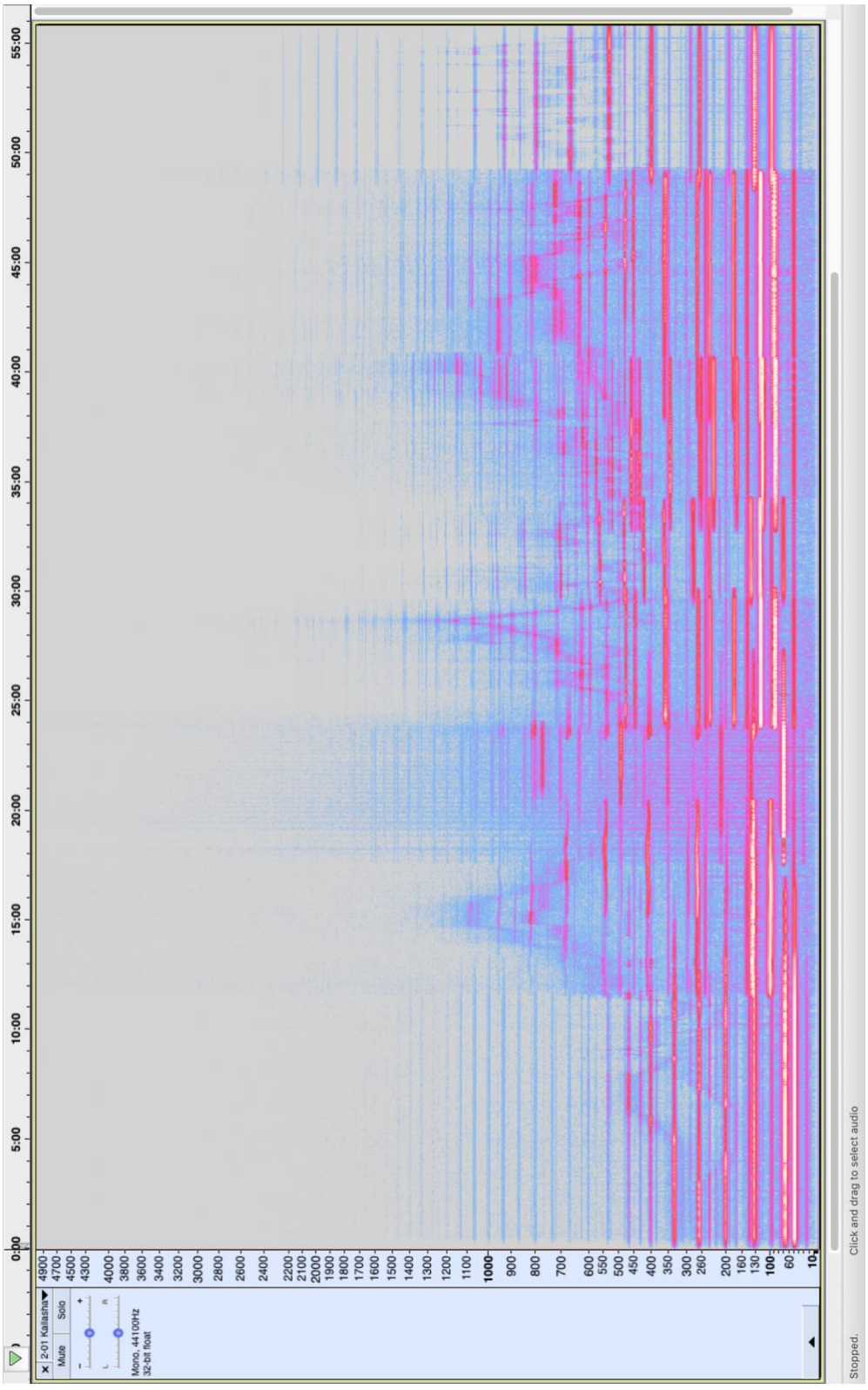

Figure 18. Spectrogram of "Kailasha." Time is shown on the x-axis, frequency is shown on the y-axis (mel scale).

5.2.2 Analysis

Radigue is not only the composer and performer of her music, but also the recording engineer, layering the tracks she produces with the ARP 2500 synthesizer on top of one another in her studio.[204] Her works are recorded to tape, which is played back during concert performances.[205] Just as in concerts featuring live performances, the placement of sound sources (here, loudspeakers; in a traditional concert, musicians) takes into account the reverberant characteristics of the performance space.[206]

"Kailasha" is part II of Radigue's *Trilogie de la mort*, a programmatic work that reflects Radigue's study of Buddhism in the years prior to its composition.[207] The liner notes contain the following (unattributed—though likely Radigue's) explanation of the title:

> "KAILASHA" is a reference to an experience drawn from real life, being transposed into an imaginary journey around the most sacred of the Himalayan mountains—Mount Kailash—considered as a path to other spheres of existence.

The circumambulation of the remote peak Mount Kailash in western Tibet is a pilgrimage undertaken by several thousand Buddhists each year. How the music relates to this program is not made explicit by Radigue. Although there are sounds in the music quite similar to the sound of rushing wind (for example at approximately 25:00), there is

204 Radigue is not, however, responsible for the remastering of her works for compact disc. In the liner notes, she cites the difficulties she and her engineer encountered digitizing her tapes, specifically "Kailasha." Artefacts of this process are audible when listening to certain deep bass tones, which can sound choppy and rattling rather than smooth and resonant.

205 "The final product is always recorded. I've always presented it as such during my concerts." Quoted in Paul Schütze, "Surround Sound," *Frieze*, 2011, https://frieze.com/article/surround-sound.

206 Since 2001, in a notable departure from her earlier work, Radigue has composed primarily for acoustic instruments, for performance in a concert setting by other musicians.

207 For example, "Kyema," part I of the *Trilogie*, consists of six sections named after "bardos" described in the Tibetan Book of the Dead—states associated with Buddhist ideas of death and rebirth.

otherwise little easily discernible programmatic content. (The use of sustained tones, however, along with their resonant qualities, and the overall pacing of the piece, does give the music a superficial resemblance to Tibetan Buddhist chant, which may conceivably be a sound associated with the pilgrimage.)

"Kailasha" begins with a thirty-second fade-in of a low, warm, gently pulsing electronic tone, and for the first eleven minutes the piece exhibits very little change at all. Every minute or two there is a slight adjustment of the filters on the sounds that cause slow and subtle changes in the audible resonant frequencies, but otherwise, everything is held in stasis. This first eleven-minute section is essentially a prelude, introducing the listener to the tones, timbres, and pacing of the work as a whole. At around 11:40 another, lower oscillating tone enters, changing the overall color of the sound, though the other elements continue unaffected. From this point on, comparatively more rapid and noticeable changes occur every three to five minutes. While slight timbral variations are still detectable in the various strata of the texture, these more dramatic shifts serve as transitions that stake out relatively stable regions in the work as a whole.

Once the music at the beginning of "Kailasha" has reached its full volume and has continued a few moments, it becomes apparent that this opening sound is not just a single, complex tone, but is comprised of at least two parts: a subtly oscillating sound pitched around D♭3, accompanied by a lower one pulsing at a slower rate an octave below. In fact, the slow pulse of the lower sound is actually comprised of a pair of alternating tones with periods of approximately two and three seconds, respectively. The higher sound, by contrast, is more like a tremolo, with a cycle approximately one second long. The interaction of these repeating cycles generates sonic patterns that continuously vary. These

oscillations are clearly present throughout the piece,; they provide a pulse that manages to propel the music in lieu of any harmonic progression.[208]

These first moments constitute the most static portion of the piece, with no other change for nearly three minutes. Only at 2:56 is there any alteration of the sound besides the patterns made by the dual oscillating cycles: a higher pitched resonance emerges, not as a separate tone exactly, but rather as the highlighting of a particular harmonic frequency. Such highlighted harmonics become more and more prevalent as the piece continues, eventually coalescing into small melodic fragments.

Manipulating the filters on the ARP 2500 to select discrete pitches is a technique Radigue applies throughout "Kailasha." As the filters can be adjusted to isolate overtones, melodic intervals of a perfect fourth, fifth, and octave are common, but Radigue also manages to produce major and minor thirds, major seconds, and the occasional semitonal and microtonal melodic interval. At times, these notes outline simple chords, for instance, a D♭ major triad in the beginning section and later, a B♭ minor seventh chord between 30:00 and 33:00. Sections where the individual tones are more easily heard as melodic occur around 35:00–38:00 and again in the closing section. At these points the sequence of pitches falls within a narrow, singable ambitus, featuring rudimentary rhythms and modest motivic repetition.

The transitions, though gradual (their beginnings are often obscured and are not always noticed until the transition is well underway), exhibit a rate of change that makes

[208] The tremolos I refer to here are generated through the use of the ARP's low frequency oscillator. However, as will be discussed below, tremolos are also evident in the music that are produced not by the instrument but by the interference of waveforms as the music is played back and reflected off walls and surfaces of the listening area.

them stand out against the surrounding, more stable sections. Often, these transitions occur through the addition or subtraction of lower frequency sounds that can be felt as well as heard and thus have the ability to affect us on a more visceral level. The transitions also facilitate identification of some of the separate strands that make up the complex texture. In a piece such as this, it is not always easy to differentiate entirely between texture and timbre. What might seem at one moment to be an integrated sound may reveal itself to be made up of multiple layers when one of those layers begins to change. (Recall the discussion of Leydon's article in the previous chapter, in which she describes how lines can blend together in performance as if they were fundamental and harmonic of a single instrument.) Through her configuration of materials, Radigue inverts the model of musical structure found in much traditional analysis of common practice works. Rather than considering rhythm and melody as a scaffold upon which timbral qualities are overlaid, this music is fundamentally timbral, with its pulsing and melodic components overlaid as coloristic qualities.

Just after the forty-nine-minute mark there is a relatively rapid (perhaps 20 seconds: between 49'05" – 49'25") and more pronounced transition into a lengthy, subdued section that closes the piece. Radigue has said that she composes by layering multiple tracks onto a single tape, each one performed by herself, added one at a time.[209] In this closing section, the layers have been thinned out, so the resulting tone sounds much less active, characterized by fewer oscillations that beat against one another at different rates. The ending thus restores a relative purity of tone, or at least, after the preceding

209 Schütze, "Surround Sound."

passage, one that initially sounds much simpler. Though, as with all of Radigue's tones, its rich, undulating quality soon becomes apparent.

One of the most striking features of "Kailasha" is that even though the signal that constitutes the piece is fixed on tape, the music produced from it is highly variable. Not only can timbral characteristics vary depending on the position of the listener in the listening space—a feature of all music, and of all sound for that matter—but individual components of the texture can recede into the background, loom into prominence, or just disappear. This is not the type of variable sound quality that occurs in a listening situation where direct and reverberant sound mesh together throughout the listening area. Because Radigue's tape music is designed to be diffuse, there is no ideal listening position, and thus, no one way in which the sounds emitted from the loudspeakers are supposed to combine into the sounds heard by the listener. "Kailasha," by calling upon us to use our bodily capabilities for engaging with the music, is able to point us toward its own affordances for action, which in turn allow us new ways of interacting not only with musical sound, but with our own surroundings as well.

5.3 Sound Waves and Texture

If we were to consider an organism such as a bat that listens as a means of echo-location, then indeed, listening would be a mode of ecological perception analogous to vision in Gibson's framework. But we do not rely on the qualities of a sound and its reverberant properties solely to reveal the existence of surfaces in the environment. In general, when analyzing music—and often when simply listening to it—such cues are

routinely factored out. For example, we note how a mass might have been composed as a site-specific work, designed to take full advantage of the reverberant properties of the cathedral as performance space. At a concert performance of the mass, a sound mixer may mix reverb into the channels of sound projected into the hall to supplement what might be perceived as a deficiency in the venue's acoustics. Such an action is concerned with the presentation of a work and is extraneous to the composition itself.

As noted in the previous chapter, Nina Sun Eidsheim explains how the extra-notational acoustic dimensions of music—"which may be simply described as the length of the reverb and the sense of clarity"—play an important role the activity of listening.[210] We don't necessarily listen for reverberation, yet if the reverberation times are outside the norm, the music sounds "wrong."[211] Thus, these compositionally unspecified features of the music's sound are crucial to our ability to hear music—to attune ourselves to the elements of musical expression in a work. (A similar act of discernment was necessary to hear music recorded on early phonograph cylinders. Listeners became accustomed to listening *through* the pops and cracks and the limited dynamic range and frequency response to nonetheless hear "natural" sound.) When acoustic conventions are broken, we become conscious of them, leading us to attribute meaning to the non-normative elements.[212]

Like light in the air, we normally consider sound to be exceptionally uniform in its distribution throughout a given listening area, reliably decreasing in amplitude as it travels according to the inverse square law. However, when we encounter a sound that appears or disappears with a slight movement of the head, or a step in one direction or another—the

210 Eidsheim, *Sensing Sound*, 61.
211 Eidsheim, 69.
212 Eidsheim, 61.

effect can be surreal and disorienting. Not only are the ordinary listening conventions of concert music broken, but the rapid change in the music challenges our understanding of how sounds behave and even how we hear. Such music forges a link between movement and listening, soliciting a participatory, gestural, and peripatetic mode of engagement.

These unexpected alterations of musical tones are due to the physics of sound, specifically, the behavior of sound waves and how they interact with one another. Sound waves are propagated through a medium, reflected off surfaces, and can combine with one another to form more complex waveforms. The sound wave, a pulse of energy rather than a substance in itself, requires a material through which to move.[213] Typically, that material is air; air molecules comprise the sum total of sound's materiality. Yet this material—the air and its molecules—is not detectable by the unaided senses. We are conditioned, therefore, to consider air as empty space. Sound, then, is perceived like wind: a series of pressure gradients, small ones sensed by our ears and larger ones sensed by our skin (and possibly our entire bodies, once the vestibular system is engaged and our balance is affected). Yet the material character of this atmosphere can, in fact, be revealed to us, perhaps not as molecules, but as a material nonetheless through its very malleability.

We continually make small, unconscious adjustments to tune into what we are listening to.[214] In a piece like "Kailasha," where the additive effects of reflected waveforms are made audible, all it takes is one of these barely conscious movements, taken in an effort of attunement, to make us aware of the discontinuous sounds in the music. Ordinarily, it is

213 "Acoustic disturbances propagate in the form of waves. A wave in a material medium may be defined as a process by means of which a disturbance from equilibrium is transported through the medium without net transport of mass." Frank Fahy, *Foundations of Engineering Acoustics* (Oxford: Elsevier Science & Technology, 2000), 7.

214 See note 168, page 17.

not our movement that will make a steady pitch oscillate as a tremolo, even less, that our movement will cause pitches to drop out of the texture. However, interference patterns of very simple sound waves—where peaks and troughs reinforce or cancel each other out—are ones that our ears can readily detect, and can result in just such musical effects.

These effects are surprising because they contradict the environment model proposed by Gibson. According to him, an enclosed medium can be "filled" with light, with sound, and even with odor. He then states that "any point in the medium is a possible point of observation for any observer who can look, listen, or sniff. And these points of observation are continuously connected to one another by paths of possible locomotion."[215] When continuity is broken, however, as in these interference patterns, and locomotion results in discretely revealed objects, we break with the world described by Gibson. His "medium" cannot account for the discontinuities of the sort encountered in Radigue's music. The sounds do not behave as we would expect the sound from a fixed source to behave. The medium is not filled with sound, in the sense of an environment illuminated by the "unlimited scattering" of radiated light.[216]

As detailed in the previous chapter, our ears are designed to deflect sound waves and redirect them into the ear. Though not moveable themselves as in other species, our ears are attached to our heads and thus to our bodies and their position relative to incoming wave fronts can be calibrated to enhance particular sounds over others. Radigue's music is composed of simple waveforms that interact with one another on an audible level. More complex waveforms, with more information, exhibit, perhaps

215 Gibson, *The Ecological Approach to Visual Perception*, 13.
216 Gibson, 44.

paradoxically, increased uniformity on a larger scale. Head movements reveal less about these complex waveforms themselves than about their sources. What we notice is not a change in the sound we hear, but a change in the sound's distance and direction from where we are located.

An analogy is possible here with one of Gibson's vision experiments, where he found that a series of alternating discs, when sufficiently dense, coalesced to form a continuous surface. He showed subjects a series of thin, plastic panels with a cut-out circle, spaced apart, arranged one behind the other and alternating black and white, as in the illustration below. When viewing thirty-six panels aligned at equal distances, the alternating black and white circles outlined by the panels coalesced into a smooth surface, forming what looked like a striped tunnel.[217] When there were substantially fewer of these of panels, no more than a dozen or so, most subjects saw individual rings with space in between them. There seems to be a certain minimum density that is required for us to perceive objects on a larger scale.

[217] Gibson, 155.

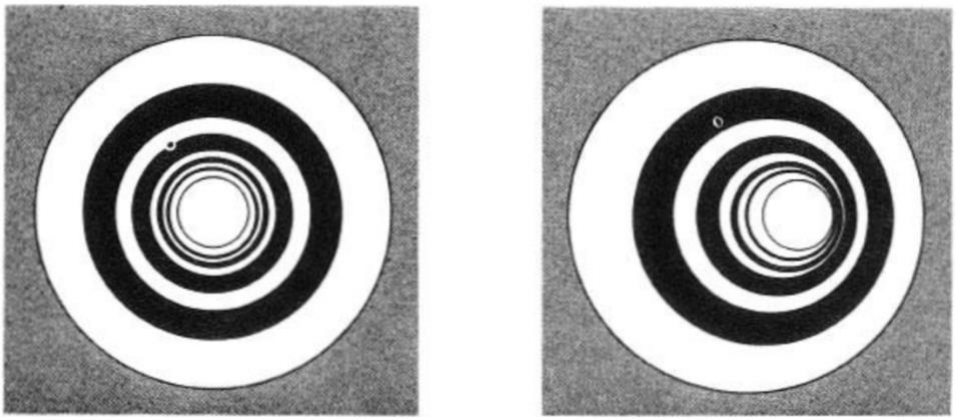

FIGURE 9.1 The optic array coming to the eye from the optical tunnel.

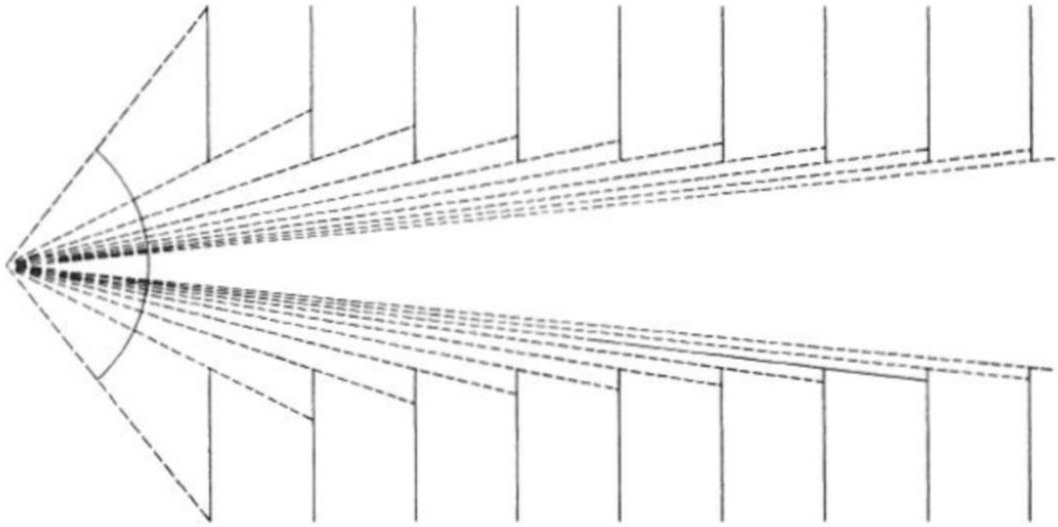

FIGURE 9.2 A longitudinal section of the optical tunnel shown in Figure 9.1.

Figure 19. Gibson's diagrams of his optical tunnel. The Ecological Approach to Perception, p. 146.

In the same manner, complex waveforms manage to suffuse a space much more densely than simpler ones, and the additive interference of individual frequencies are not nearly as apparent. In his experiment, Gibson found that density led to a "surfaciness" that

obscured the composite structure of the tunnel. In regard to sound, density similarly obscures the composite waveforms that constitute its structure. Electronic music of the variety composed by Radigue, however, can reveal this composite wave structure and turn it toward productive aesthetic purposes.

As Moylan states, the perception of a sound's distance is a function of that sound's timbral detail. Timbre, by definition, is the distribution and relative strength of partial frequencies and this spectral information is a key component in the auditory system's ability to identify where a sound is coming from. An analog subtractive synthesizer—such as Radigue's ARP 2500—has a limited ability to recreate the full spectrum of harmonic and inharmonic vibrations that comprise any instrumental timbre. As a result, there is less detail—less high frequency information—that will rapidly dissipate with distance. These waves, then, with their reduced timbre, are not far removed from sine waves. And since these are simple waveforms, they combine with their reflections to form waves of similar shape but varying intensity. Thus, not only are these sounds without timbre, they are sounds without reverb: reflected waves combine with the direct signal, either attenuating or reinforcing it. The sounds convey no information about their source and its location, nor about the characteristics of the space (actual or engineered) in which they were recorded. Shorn of these characteristics, the sounds cannot but be perceived as occurring where they are heard.

Matthew Nudds (2008) explains this important distinction between the location of sounds and the locations of their sources, and of our ability to recognize the difference.[218]

218 Matthew Nudds, "Sounds and Space," in *Sounds and Perception: New Philosophical Essays*, ed. Matthew Nudds and Casey O'Callaghan (New York: Oxford University Press, 2009).

To determine the location of a sound source, the auditory system groups together frequency components that likely belong to a single source. For example, a set of frequencies that belong to the same harmonic series, or that start and end at the same time, would be attributed to the same vibrating object. By contrast, Nudds explains that "the sounds that we hear are instantiated where we are." By this he means that even though we perceive sounds as being connected with their presumed source, the physical sound waves we hear, the changes in pressure gradients at the eardrum, exist in the same place that we happen to be. This separability between the sound and its source allows us to hear sound sources as remaining constant even as the qualities of the sound may change. To illustrate, Nudds uses the example of pulling a hat down over the ears, muffling any incoming sound: we don't believe that the source of the sound has changed, even though the sound we hear has been altered.[219] When "Kailasha" is played, its sound does not disperse uniformly throughout the environment. Therefore, we cannot attribute the sound we hear to a fixed source. The ambient aural array that reaches our ears is so drastically distorted that the medium itself is revealed as irregular, as having its own texture.

Marratto writes that "textures cannot show themselves otherwise than to a body capable of exploratory movements."[220] Just as with the texture of a fabric, the texturing of space with sound requires a gesture in order to be perceived. We don't simply touch a fabric's texture, we feel it, which means running our fingers over it. And just as the things we touch lie where we touch them, the sounds we hear are located right where we are, not

219 Nudds, 93. Similarly, Moylan claims that the primary cue for distance of a sound source is "the loss of low-amplitude (usually high-frequency) partials from the sound's spectrum with increasing distance," i.e., the further away a sound source is, the less timbral detail it has.) Moylan, *Understanding and Crafting the Mix*, 2007, 28.

220 Marratto, *The Intercorporeal Self*, 28.

at a distance beyond our peripersonal space. Likewise, with Radigue's music, its sonic texture is not located at a distance, as in Gibson's description of viewing a landscape, but it *is* where we hear it. We thus have an awareness of space not in terms of direction, but as a depth revealed through the texture provided by the music. As "Kailasha" is distributed through the environment unevenly, it carries with it its own palpable substance. Its texture does not reveal the materiality of the surfaces that reflect it, but instead reveals the music's own peculiar materiality. It is a dramatic expression of music's ability to have presence.

Note that this differs from the concept of acousmatic sound, a sound whose source remains hidden from view. Acousmatic sound entails a multimodal conception of listening in which both hearing and vision take part. Its distinctive character is a result of severing the connection between sound and visible source. Acousmatic sound does have a source; it is just that for the listener that source is not reconcilable with corroborative visual perceptions. Sound phenomena that result from the interference patterns of simple waveforms, however, do not have a hidden source: they appear to us as having no source at all. They are atmospheric.

5.3 Space, Gesture, and Atmosphere

Radigue's music might be said to achieve a heightened sense of spatiality precisely because it does not rely on the representation of sound source locations. "Kailasha" creates an impression of space that is not so much filled with sounds but that is *composed of* sound. The awareness of physical space arises in conjunction with the perception of the musical sounds in the activity of listening. By considering listening as an embodied activity, this

simultaneous appearance of space and sound makes sense. An awareness of space can only arise with respect to possibilities for bodily motion—and if music can solicit our attention, and we then take up the affordances it offers, then music can shape our spatial experiences.

Composer, work, listening subject, and analytical object: Radigue's music forces us to consider each of these by foregrounding our movement through the spaces we inhabit alongside the music. This entanglement of composition, performance, and listening helps explain why such an ordinary concept as space can so easily slip from our grasp. Although we readily admit that there is always an interaction between ourselves and the things we study, we consistently assume that there is nothing that keeps us from stepping back out of the equation, factoring out the subjective to leave the objective. But maintaining the separation between objective music and subjective listener, however much they intermingle, frustrates an account of how music can motivate and inspire new actions and, in turn, new meanings.

The inexhaustibility of the meaning of the work of art, according to Merleau-Ponty, lies in its deployment of a mutually inhabited world where there are always new perspectives to be taken, and where these new perspectives, in turn, give us a new world "to express and think about that envelops and exceeds those perspectives."[221] Merleau-Ponty, discussing the aesthetic significance of a line drawn on a piece of paper, writes that the viewer is invited to "take up the gesture which created it."[222] As he makes this claim in the context of an inquiry into interpretation of abstract or expressionist art, it applies well

221 Maurice Merleau-Ponty, "Indirect Language and the Voices of Silence," in *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, ed. Galen Johnson, trans. Michael Smith (Evanston, IL: Northwestern University Press, 1993), 89.
222 Merleau-Ponty, 88.

to the non-representational aspect of music. Such "taking up" does not suggest mere repetition or reenactment of a gesture, but an engagement with it, as in a game, where one gesture is responded to by another's gesture, co-creating the interaction that *is* the game. This is why a communicative model of the transmission of meaning, whether affective or narrative, is not always appropriate for music. The fact that different listeners understand different things, and that new meanings can disclose themselves over time, means that understanding involves more than the recognition of musical gestures. Instead, it is our own complementary, corporeal engagement that is solicited by the musical gesture, an engagement which, in turn, affects our manner of inherence in a world with depth.

Radigue's music is an illustration of Gernot Böhme's concept of atmosphere. For Böhme, atmosphere is "the relation between environmental qualities and human states."[223] With respect to its intermediate status, he writes, "atmosphere is the common reality of the perceiver and the perceived. It is the reality of the perceived as the sphere of its presence and the reality of the perceiver, insofar as in sensing the atmosphere s/he is bodily present in a certain way."[224] We shall have an opportunity to discuss this concept at length in the following chapter, but here I will provide an introduction to the concept and how it arises out of a discussion of "Kailasha."

If I and a companion stand at one end of a long, high, dim, gothic cathedral and gaze back toward the nave, we may sense a large space surrounding us. If I then proceed to light a candle and to hold it between us, the space we sense is no longer vast, but close and intimate: the space is modulated by the candlelight, whose illumination does not reach very

223 Böhme, *Atmospheric Architectures*, 12.

224 Gernot Böhme, "Atmosphere as the Fundamental Concept of a New Aesthetics," *Thesis Eleven* 36, no. 1 (August 1, 1993): 122, https://doi.org/10.1177/072551369303600107.

far beyond our own peripersonal space. Our space of action is now circumscribed and we can say that the *atmosphere* created by lighting the candle is largely responsible for our sense of a small space—in spite of our knowledge of the dimensions of the cathedral in which we are standing.[225] There is nothing metaphorical here, nor is this an optical illusion of some sort. We sense the space in which we are immediately engaged as an intimate one.[226] The candle *modulates* the space, to use psychologist Christoph Michels' term, which in turn affects the atmosphere in which we are situated.[227]

This modulation shapes our perception by guiding our attention and making us aware of particular characteristics of our surroundings.[228] Radigue's sonic modulations of the places in which her music is heard are akin to the modulations that can be accomplished through lighting. Thus, it is not simply a figure of speech to say music is atmospheric, for it is a major contributor to atmosphere as defined in Böhme. Neither is the space created by music illusory or metaphorical. Our perception of space is modulated through the atmosphere we find ourselves in.

Böhme proposes atmosphere as a way to understand how an object and subject relate to one another when they are copresent in an environment. The aesthetic vocabulary Böhme uses thus recalls Merleau-Ponty's idea of "compossibility," first discussed above in Chapter 1. We can then ask about things that share the same space and how we manage to

225 Merleau-Ponty writes that when he casts his gaze at a portion of the landscape, "other objects recede into the margins and become dormant, but they do not cease to be there." Merleau-Ponty, *Phenomenology of Perception*, 2012, 70.

226 Such a space might be perceived differently were we alone in the cathedral at night with only a candle. Perhaps we might have a sense once again of the cathedral's grand scale and its vast enclosed space. There are a host of factors at play as we reckon with our surroundings.

227 Christoph Michels, "Researching Affective Atmospheres," *Geographica Helvetica* 70, no. 4 (October 5, 2015): 255–63, https://doi.org/10.5194/gh-70-255-2015.

228 Michels.

accommodate those things that seem incompossible. The answer, according to the writers we have been discussing here, is that is we accommodate them spatially, for we always achieve understanding through a mobile body. As Marratto writes:

> It is only for a motor subject that objects have a past and a future, a surface, a *texture*, and an inside. This is because the embodied subject … circumspectively absorbed in the world, lets the things in its environment show themselves according to possibilities determined in reference to itself as a bodily 'I can.'[229]

Or put another way, Radigue herself says, "in order to avoid being annoyed by any type of sound, you just have to make some music out of it … through your ears, through the way of listening."[230]

Böhme says that when we look out at the world, our eyes roam from one object to another, visually exploring our surroundings in a manner that generates the depth in the space we inhabit. It stands to reason, then, that our space is indeed transfigured when confronted with the incompossible, for the incompossible introduces an imbalance into our "maximal view" of the environment and requires a reestablishment of equilibrium.[231] If we are to understand the novel thing as something other than imaginary, then we need to engage with it, to take up a perspective that will allow us to encompass what is unfamiliar and make it meaningful for ourselves. Aesthetically, we do this through listening to the affordances the music offers and responding to the gestures it solicits. To understand "Kailasha," then, we must take a perspective in which space is *not* empty, and in which music can reveal possibilities for action we may not have imagined.

229 Marratto, *The Intercorporeal Self*, 122. Emphasis added.

230 Max Dax, "Eliane Radigue: An Interview," *Telekom Electronic Beats* (blog), October 11, 2012, https://www.electronicbeats.net/eliane-radigue-an-interview/.

231 See Part I, p. 63: Taylor Carman on Merleau-Ponty and "optimal grip."

Chapter 6: Steve Reich, *Electric Counterpoint*, Movement I, "Fast"

6.1 Introduction

In this dissertation, I have considered a variety of musical factors that contribute to our perception of space. To do so, I have drawn upon a tradition of phenomenological analysis of music wherein I, as a theorist, through an act of introspection reflect upon my musical experience and attempt to convey that experience to a reader. This approach is warranted by music's undeniable affective force—a force that can only be ascertained by methods which take into account the effects the music generates for a listener. I have combined this phenomenological approach with an aesthetics of sound, that is, an aesthetics that considers musical phenomena a subset of sonic phenomena. In addition, I considered fixed media works alongside commercial recordings of notated compositions as comparable in terms of how they are realized through performance. This required the introduction of both acoustics and media theory to help explain the impact recording technology has had on our musicking.

Part I considered the concept of space that is often at play in contemporary composition and analysis. By going back to the early days of sound recording, I showed how an understanding of space concerned with the location of sounding objects was instantiated in both the musical practices and in the discourse surrounding the emerging technology. But this directional aspect of musical sounds is just one type of space. Music

can also manipulate our sense of the volume and the character of the space we inhabit—what Böhme refers to as "atmosphere."

In Part II, I began by elaborating the notion of phenomenological space with respect to music listening and explained how it emerges out of Maurice Merleau-Ponty's theory of embodied perception. Phenomenological space, I contend, encompasses a much broader range of spatial experience than the location-based space generally evoked by compositions labelled "spatial." Drawing on the work of Merleau-Ponty, James J. Gibson, Scott Marratto, and others, I deploy an affective model of spatial perception—a model that links perception with the moment to moment needs and desires of the perceiver—to explain how musical modulations of space occur. I began by defining spatial qualities through a pair of simple binary descriptors: near and far, and large and small. These are relative terms, and importantly, relative to situated listeners, each with their own singular perspective.

The analyses in part two demonstrate some of the ways in which spatial perception is altered by the music we hear. The space we inhabit can suddenly seem to expand or contract as we engage with different musical timbres and textures. Because this engagement is affective, that is, in accordance with the listener's needs and desires, space is accessed affectively. Its affective qualities are what allow us to perceive space at all. Thus, when a space seems to us as large and small, near and far, full and empty, we are engaging with its affective character rather than its actual physical dimensions.

In this chapter, I will draw upon the concepts introduced in my earlier analyses to provide an account of the first movement of Steve Reich's *Electronic Counterpoint* (1987). In particular, I will use them to offer a hermeneutic analysis of the spatial experience

afforded by a moment in the first movement, thereby modeling a phenomenological method for grounding interpretation in embodied, rather than strictly discursive, practices. As noted above, the particular recording we use plays a role in the spatial aspects of a piece that we cannot ignore; therefore, the following analysis refers to the 1987 Nonesuch recording of the piece performed by Pat Metheny. [232] This recording provides an example of how the analytical category of texture can provide opportunities for complex spatial experiences. Such experiences demonstrate that music's capacity for modulating our spatial perceptions is a significant source of its affective power.

It is also important to clarify the concept of texture at work in this chapter as compared to that used in the analysis of "Kailasha." In the previous chapter I discussed the texture of space and described how it is not simply a metaphorical concept, but an extension of those surface textures we perceive haptically: as with a textile fabric, we perceive spatial texture through physical gestures—in this case, through those gestures afforded by the music. In this chapter, however, I will be turning back to the traditional music-theoretical concept of texture.

6.2 Musical Texture

Texture remains a loosely defined parameter in music theory, functioning similarly to genre as a means of taxonomic classification.[233] Interestingly, outside of anglophone

232 Steve Reich, "Electric Counterpoint," *Different Trains/Electric Counterpoint*, Elektra/Nonesuch CD 7559791762.

233 Jonathan Dunsby finds its treatment in *The New Grove* (1980) as "lacking conviction." Indeed, the entry in the Grove's online resource can be frustratingly vague, as when it lists George Crumb as a creator of "distinctive" textures, but does not explain what those textures are or why they are distinct. Jonathan Dunsby, "Considerations of Texture," *Music & Letters* 70, no. 1 (1989): 13. "Texture | Grove Music," accessed June 18,

musicological discourse this music-theoretical term seems to have no equivalent. Jonathan Dunsby mentions the word *Klangstruktur* and *Grove Music Online* provides *Satz* as comparable terms in German, but neither conveys the specific sense of the English word—both are more general, and the former is rarely used.

The textbook varieties of texture are monophony, homophony and polyphony (occasionally supplemented by heterophony). Broadly speaking, these textural types can be understood as the number of individual instrumental lines in a piece and their interrelationships. In *Structural Functions in Music*, Wallace Berry examines these traditional categories and reveals their composite nature.[234] He enumerates the elements that distinguish the traditional texture types and groups them into qualitative and quantitative features. Qualitative features concern the degree of uniformity or independence between melodic lines with regard to motion and rhythm. The quantitative element is simply the number of instrumental and vocal lines that comprise any particular musical passage, including their density as defined in terms of proximity in pitch. When lines are sufficiently dense, Berry notes how they merge and take on a background function.[235] With this observation, Berry introduces a Gestalt interpretation of musical texture (one that Leonard Meyer adopts as well, in his discussion of music's Gestalt—and non-Gestalt—characteristics). That is, beyond a quantitative account of the number of

2019, https://www-oxfordmusiconline-com.ezproxy.cul.columbia.edu/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000027758.

234 Wallace Berry, *Structural Functions in Music* (New York: Dover, 1987), 193–95.

235 Berry, 249.

voices, there is a qualitative interpretation that imposes a figure-ground distinction, and thus a hierarchy, among the voices.[236]

Berry also refers to "musical space," which he defines as the ambitus that defines the boundaries or compositional limits of the piece. Musical space is a range over which a particular musical parameter operates, i.e., space as the set of values between two limit points. It is thus a term of linear measurement and is separate from the idea of space investigated in the second part of the present study, which is predicated upon the perception of depth.

Pablo Fessel approaches texture from a listener's perspective and describes how individual lines can meld and form similarly distinct "autonomous planes," or strata.[237] Strata are useful for describing the music of Steve Reich. While Reich's title suggests that *Electric Counterpoint* is a contrapuntal composition, the strata that emerge from the grouping of the parts exhibit qualitative features as well. Fessel's structuralist approach is listener-based: if a particular voice or layer cannot be distinguished by a listener, then Fessel does not consider it as a textural stratum.[238]

Following Berry, I will consider texture in terms of the number of lines and their relationships with one another, while making use of Fessel's functional concept of musical strata and their roles relating to background and foreground structures.

236 See, for example, Meyer's discussion of polymeter, where he writes of "the tendency to hear a relatively uniform part of a texture as being as uniform as possible so that other figures will appear more clearly shaped…." Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago: University of Chicago Press, 1956), 122.
237 Pablo Fessel, "Hacia Una Caracterización Formal Del Concepto de Textura," *Revista Del Instituto Superior de Música*, no. 5 (1996): 80.
238 Fessel, 80.

6.3 Maximal Grip

In *Electric Counterpoint*, listener position and listener movement do not have the same effect on the music as we noted in the Radigue piece: small shifts of the body do not result in large changes in its sound. Yet these motions are nonetheless ones that are taken by a perceiving body in search of what Merleau-Ponty calls a "maximal grip" on an object of interest. He likens this movement to the activity of viewing paintings in a gallery, and that not only can things seem too far away, they can also seem too close. What is sought, he writes, is "equilibrium:"

> For each object, just as for each painting in an art gallery, there is an optimal distance from which it asks to be seen—an orientation through which it presents more of itself—beneath or beyond which we merely have a confused perception due to excess or lack. Hence, we tend toward the maximum of visibility and we seek, just as when using a microscope, a better focus point, which is obtained through a certain equilibrium between the interior and the exterior horizons.[239]

This search for an orientation that provides a maximum of visibility is motivated by solicitations from the music, the "asking" that Merleau-Ponty describes. If the music seems to require a different perspective in order to hear it better, and we are affected by it and heed the call, then we are engaging with the music through our gestures, that is, through our adjustments in posture or position.

The first movement of *Electric Counterpoint* is a good case study for examining the relationship between musical texture and phenomenological space because of the clear

239 Merleau-Ponty, *Phenomenology of Perception*, 2012, 315–16. Gibson makes a similar point, using similar language: "[The Ecological approach to visual perception] begins with the flowing array of the observer who walks from one vista to another, moves around an object of interest, and can approach it for scrutiny, thus extracting the invariants that underlie the changing perspective structure. Gibson, *The Ecological Approach to Visual Perception*, 290.

division between its strata: a lower bass register that contains repeated pitches, and a rhythmically active upper register. Though these strata become more complex and less registrally distinct toward the end of the movement, their changing character and interactions are relatively easy to follow. This movement also illustrates my contention that our spatial understanding of a piece is not simply a byproduct of metaphor or analogy, but is primarily a function of how we perceive musical sounds. Musical perception emerges in the dual, mutually constitutive moment in which subject and object emerge, where we can get a grip on the music and in turn discover ourselves as being gripped *by* the music. Though this spatial relationship is not metaphorical, there is an imaginative process going on: our proprioceptive knowledge (i.e., our inherent knowledge of our body's capabilities) guides us toward taking up certain stances while dismissing others—stances that allow for an optimal grip upon the thing that interests us.

6.4 Analytic Description

Although *Electric Counterpoint* is composed as a work to be performed live (by at least one musician) for a concert audience, the recording can also lay claim to an ontological status as a musical work itself, just as Frith describes for rock music.[240] Frith, however, claims that the rock recording is ontologically prior to any other mode in which the work may exist. This is not the case for *Electric Counterpoint*, although the relatively wide distribution of the Nonesuch recording—performed by a guitarist who is a well-known recording artist—along with its subsequent sampling in electronic dance music,

240 See Part I.

shows that ontological priority is not always easy to assign. For this work, we can at least say that the recording exists alongside live performance as a manifestation of the same composition. Though it is not prior, neither is it derivative or secondary. Regardless, whether we consider a live performance of the piece or an instance of playing a recording, the experience can be framed as one of "music listening" as opposed to, say, one of "listening to a recording of music" (or to a reproduction or representation of a performance), though this latter frame is freely available to the listener as well. This enculturated practice which considers listening to recorded music *as* music listening allows us to address the recording as an expression of the work on par with live performance, or for that matter, with its inscription in a score. Thus, elements fixed in a particular recording, like choices of microphone, panning of tracks to channels, and processing effects, are part of the expression of the work as well. (In a work for electric guitar, it is not altogether clear whether reverb should be considered a studio effect or an inherent part of the timbre of the instrument: amplifiers for electric guitars typically include a reverb control.[241] Therefore, the spatial contribution of reverb in this case has to be considered in light of conventional performance technique.[242])

The score for *Electric Counterpoint* contains twelve staves with a total of thirteen parts: one staff for each of eight guitars, bracketed together, that play repeated patterns of single notes; a pair of bracketed staves for two additional guitars that strum trichords and

241 "Electric Guitar | Grove Music," accessed May 24, 2018, http://www.oxfordmusiconline.com/view/10.1093/gmo/ 9781561592630.001.0001/omo-9781561592630-e-0000008690.

242 This ability to incorporate an electronic simulation of distance into the characteristic timbre of an instrument is another example of how timbre as an instrumental quality has the capacity to convey space sonically. See previous chapter on Crumb's *Black Angels*.

dichords respectively; a single part for solo guitar labeled (tellingly) “Live”; plus one staff for two bass guitars, doubling each other at the octave.[243] (See Figure 19.) Performance instructions indicate that the piece may be played by either a guitar ensemble or a single guitarist playing the live, solo part accompanied by a taped recording of all the other parts (the latter instrumentation is specified first, lending it a degree of primacy over the guitar ensemble). The hybrid nature of the instrumentation of the work further complicates an already nebulous ontology. But that dual designation means that we do not have to go out on a limb to consider the playing of a recording to count as a musical performance rather than, say, a doubly mediated instance of a work.

[243] The label “Live” is telling as it defines role of the soloist in terms of recorded music. Recorded music is thus the unmarked variety of musical expression here (i.e., the norm), a status reinforced by Reich’s use of expression markings such as “fade in” and “fade out.”

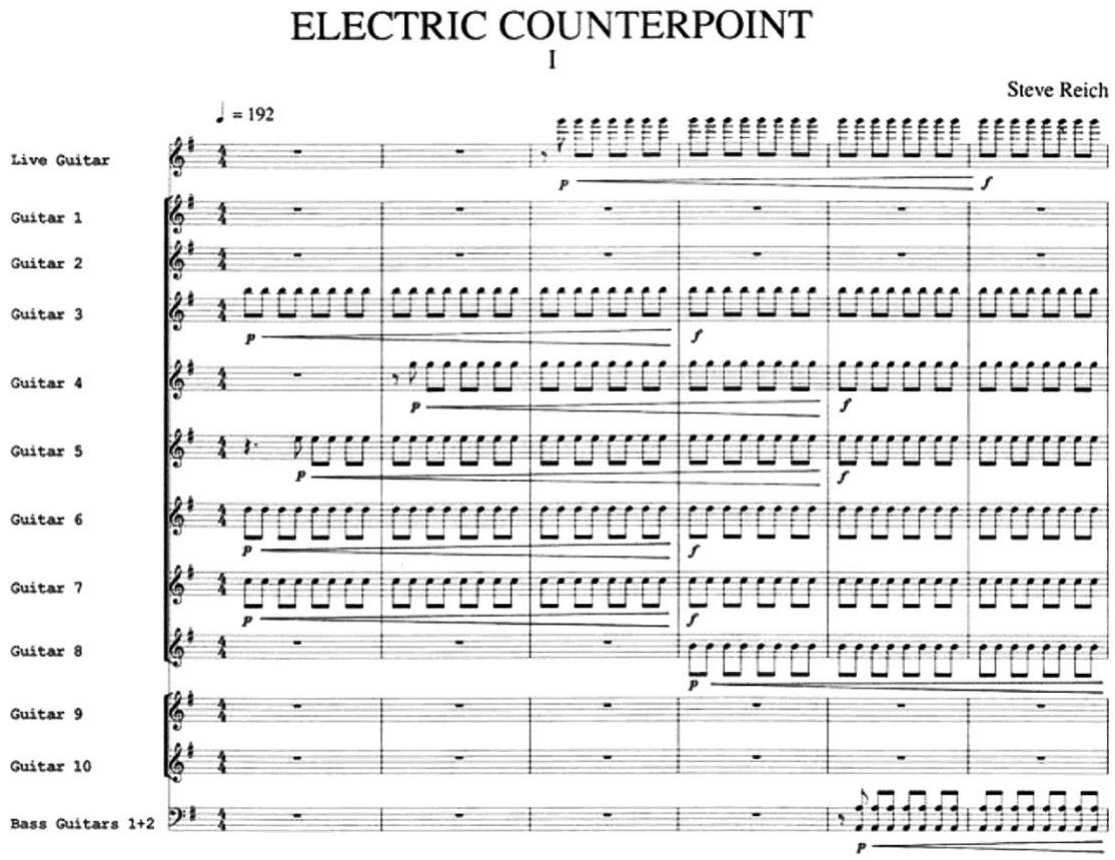

Figure 20. First page of *Electric Counterpoint*, mm. 1–6.

On the Metheny recording, when new voices are introduced, they are panned to either the left or the right of the stereo field; the lead guitar part maintains a position in the center. This assists the listener in grasping the basic cumulative process involved in the development of the hocketing pattern, as similar fragments can be heard in the guitar parts as they continue to multiply to the left, right, and center of the listener. This hocket pattern is not only key to this work, it is key to a large portion of Reich's oeuvre. Hocketing is a polyphonic technique where the same melodic material is shared between different voices, typically in cannon, with staggered soundings of the voices accomplished through the

strategic placement of rests that offset the repetitions and flow of the melodic statement. Reich's use of this technique can be traced back to his early work on magnetic tape, mixing one tape loop with another played back at a slightly different speed to generate a timbre (manifested as phasing, reverb, and echo) slowly developing into a multiline texture (as overdubbed delay) while using the same musical material.

There is nothing in the score of *Electric Counterpoint* to specify the assignment of the instrumental voices to any particular location within the stereo field; although the voices on the Nonesuch recording are assigned separate channels, the piece could well have been recorded differently. But a rationale for this choice seems to lie in the composition. Though the thirteen voices eventually sound together, they enter singly, at first playing only small fragments of melody that consist of two, three, or four eighth-notes, that then additively, note by note, become longer phrases. These staggered entrances suggest that the lines are to be understood as maintaining their status as separate streams running alongside one another, as opposed to, say, the tight mesh of Ligeti's micropolyphony. The choice to record the parts in this way, each one panned to an alternate channel after it enters, seems to be consistent with the way the movement is composed: the mixing technique reflects the texture specified in the composition.

The first section of the movement provides a series of pulsing chords in eight voices—diatonic hexachords, ambiguous with respect to an implied root. They are voiced across a wide spectrum, from E2 to E6. Processed with reverb, the multiple pitches convey an immediate fullness and resonance that gives the sense not only of the music being performed in a large room (or perhaps a small hall) but also of filling that space. The harmonies are supported by two bass guitars, doubling one another an octave apart. Each

chord sounds for eight measures before transitioning to the next one through a composed crossfade in which the voices grow softer and drop out one by one, and then, before the others have all faded out, re-enter—again one by one—with a new harmony and proceed to grow louder. The second section begins with two voices playing short, repeated melodic fragments in unison. Other voices are added individually, playing snippets of the same melody, offset horizontally from one another by one or two eighth notes and, consequently, vertically by a variety of diatonic intervals. These additions continue until there are a total of nine guitars playing the dense hocketing pattern. Once this point has been reached, the third section commences by reintroducing the pulsing hexachord pattern in guitar parts 9 and 10 and the bass guitars in octaves; the hocketing continues as before, uninterrupted.

6.5 Space and Texture

By examining the contrasts between sections, we can uncover some of the factors that structure space in this piece. The opening measures of the first section establish the spatial dimensions that are in play: the performance seems to take place in an enclosed room that is smaller than, say, an auditorium, but larger than the average living room, judging from reverberation times and the proportion of direct to reflected sound. The bass suffuses and fills the listening space, while cleanly articulated higher notes are more clearly localized as very near the listener—occupying a relatively discrete physical space, locations that could easily be pointed to. Their spatiality is topographic, or as Gibson calls it, metric: the origins of sound sources are specifiable in terms of direction and distance from the listener. The beginning of the second section (Figure 20, rehearsal number 11, measure 110; approximately 2:10 in the recording) is marked by a change in texture that

reconfigures the relationship between the instrumental sounds and the space in which they occur. The reverberations of the guitar convey the impression of a volumetric performance space of a specific size, but the thin, treble-heavy quality of the music (the bass guitars having dropped out) no longer seems to fill out that space as completely as before.

This second section of the movement begins at approximately 2:10 in the recording, as the electric basses and the guitars strumming chords fade out. Each new voice in the texture enters *piano*, doubling the lead guitar, with a crescendo to *mezzo-forte*. With each crescendo, the lead guitar itself fades out, pauses for three to three-and-a-half beats, and then begins a new pattern, again building up through an additive process. The listener is invited to follow along as the new melody is created, with the existing voices serving as a busy accompaniment. This invitation is proffered in part through dynamics, as the new patterns are introduced by the soloist at a louder dynamic level (however that dynamic level is achieved, that is, whether through the player's efforts or the engineer's). The incremental accretion of the melody, note by note, also invites the listener to follow along, as if being asked to join in a game. Attention is further hailed by the placement of the solo voice in the stereo field—in the center, with supporting voices panned to the sides.

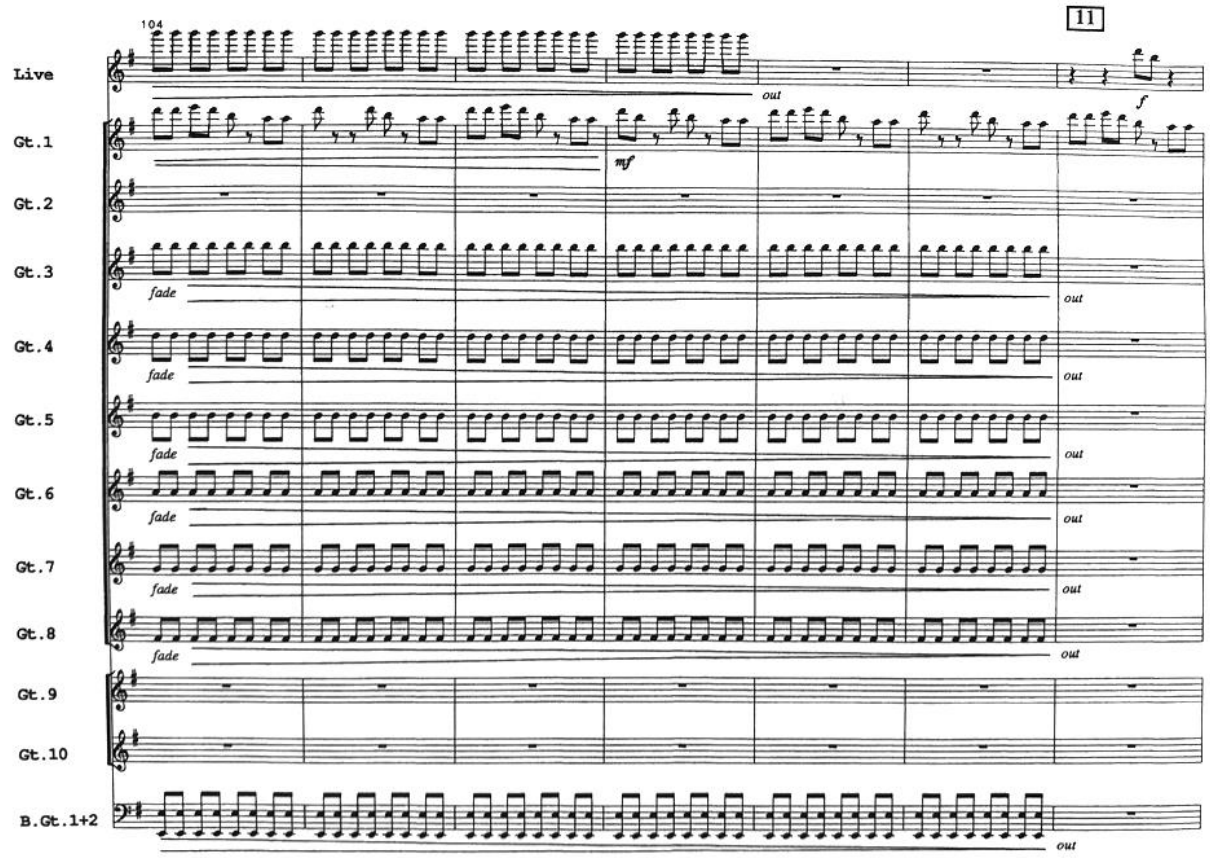

Figure 21. ***Electric Counterpoint*****, Movement I, mm. 104 – 110.**

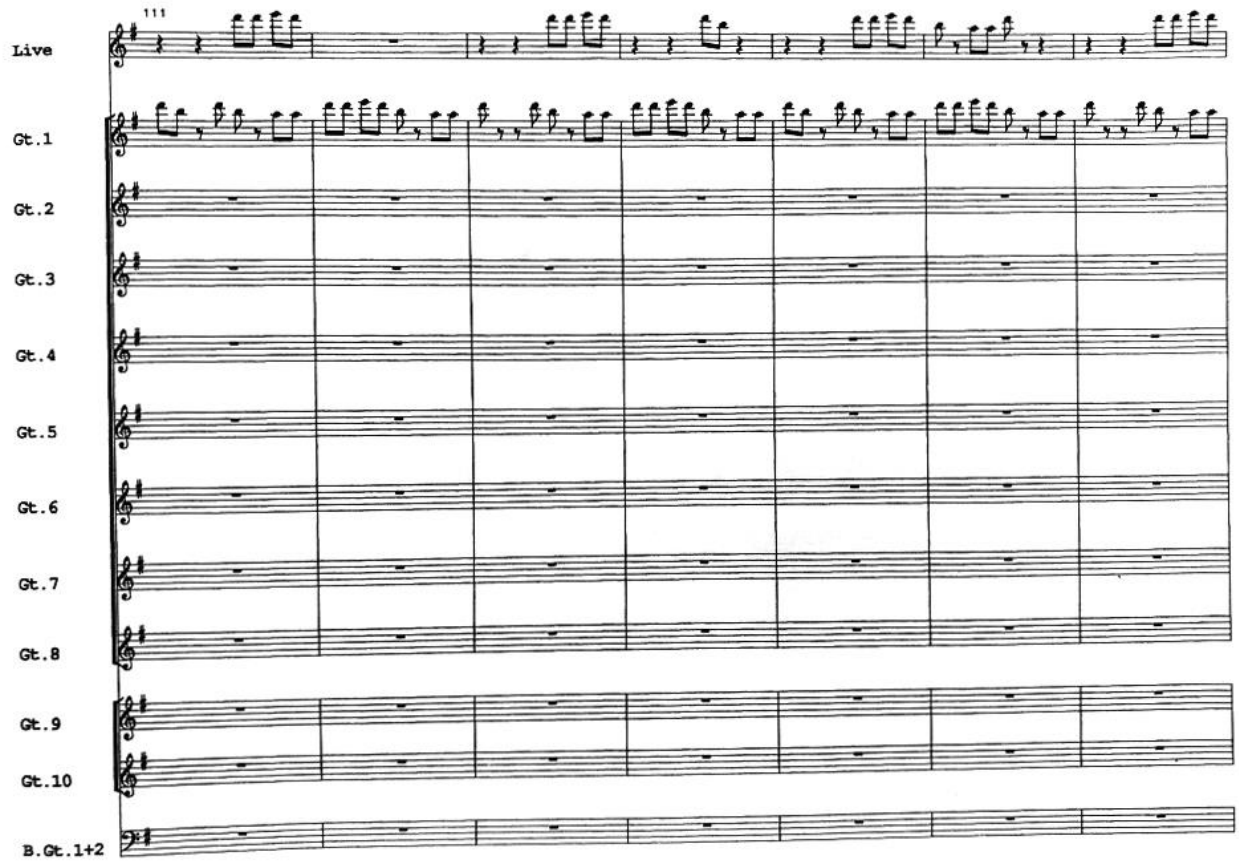

Figure 22. Electric Counterpoint, Movement I, mm. 111 – 117.

In the sudden thinning of the texture, the listener loses the sense of being awash in lush, reverberant harmonies. Instead, the listener will likely notice the stream of clear, precise attacks in the guitar part. As mentioned above, the guitar has some familiar electric-guitar style reverb here that is noticeable not just as timbre, but as a component of sound that indexes the walls of a discernible enclosure, such as a very large, empty room. But the guitar is not placed far away from the listener in the sound design of this recording: crisp articulation—the dry onset of a plucked string—provides a cue that locates the sound source quite close to the listener's body. Our ears can distinguish between direct and reflected sound, and the strength of attack on each note helps ensure we interpret the

instrumental sound source as one nearby (see discussion of Moylan and markers of distance in Part I).

The result of this particular staging of the performance configuration is that individual instrument sources appear nearby while added reverb denotes a larger, more encompassing, space—larger, yet one that remains musically empty. The music is sounding here by us, and not out there in that further space indicated by the recording. It may be that if we were located elsewhere in this alternate, projected performance space, we would hear *that* space as being full of music as well, but the space beyond these very close instruments within our reach is not the source of any additional musical objects to attend to. We can let it recede into the background. But even if that more distant space serves as a background arena for the foregrounded, and proprioceptively close, musical activity we do attend to, it does not disappear. Merleau-Ponty writes that even as other objects recede into the periphery ("*en marge*") as we focus on a particular portion of our surroundings, "they do not cease to be there."[244] They are still at our disposal, he writes, as are their respective "horizons," that is, those perspectives in which our own object of interest would become peripheral. The persistence of this horizon is crucial to how we understand the climax of the movement.

More voices enter, one by one, until nine guitar parts are playing together (Figure 22). A curious spatial effect emerges from this thicket of interlocking voices, one that is affective and contributes to the emotional content of the music. The fact that each voice is playing the same melodic material, metrically displaced from one another, tightly weaves

244 Merleau-Ponty, *Phenomenology of Perception*, 2012, 70. The translation of "*marge*" as "periphery" is Colin Smith's from the 1958 English edition; all other translations are from Donald Landes's 2013 edition. Landes uses "margins" here. The original passage is on page 82 of the 1945 French edition.

these lines together and confounds any attempt by the listener to unravel them. Moreover, the melodic material has a very narrow ambitus and contains several repeated notes. This conceals the independence of the lines and further causes them to mesh together. Berry attributes the fusing of separate lines to their density, that is, the intervallic distance between the lines. But the repeated fragments and the additional patterns that arise from simultaneous pitches or octave doubling among the different parts also contribute to the effect. Rather than an indistinct blur, what is created is something more akin to the texture of a jacquard fabric, with a varied surface that prevents the entire construction from slipping into the background. As in the moment discussed in Ockeghem's *Missa mi-mi*, a large performance space is indicated, but this time, in contrast to the Clerks' Group recording, we are not permitted access to it. The effect is a sense of being hemmed in by the music, because the space established at the beginning of the piece, and that is still indicated by the reverberant quality of the guitars, is larger than that which confines the listener, who is encircled by a net of sounds whose proximity is made evident by their clear, distinct attacks.

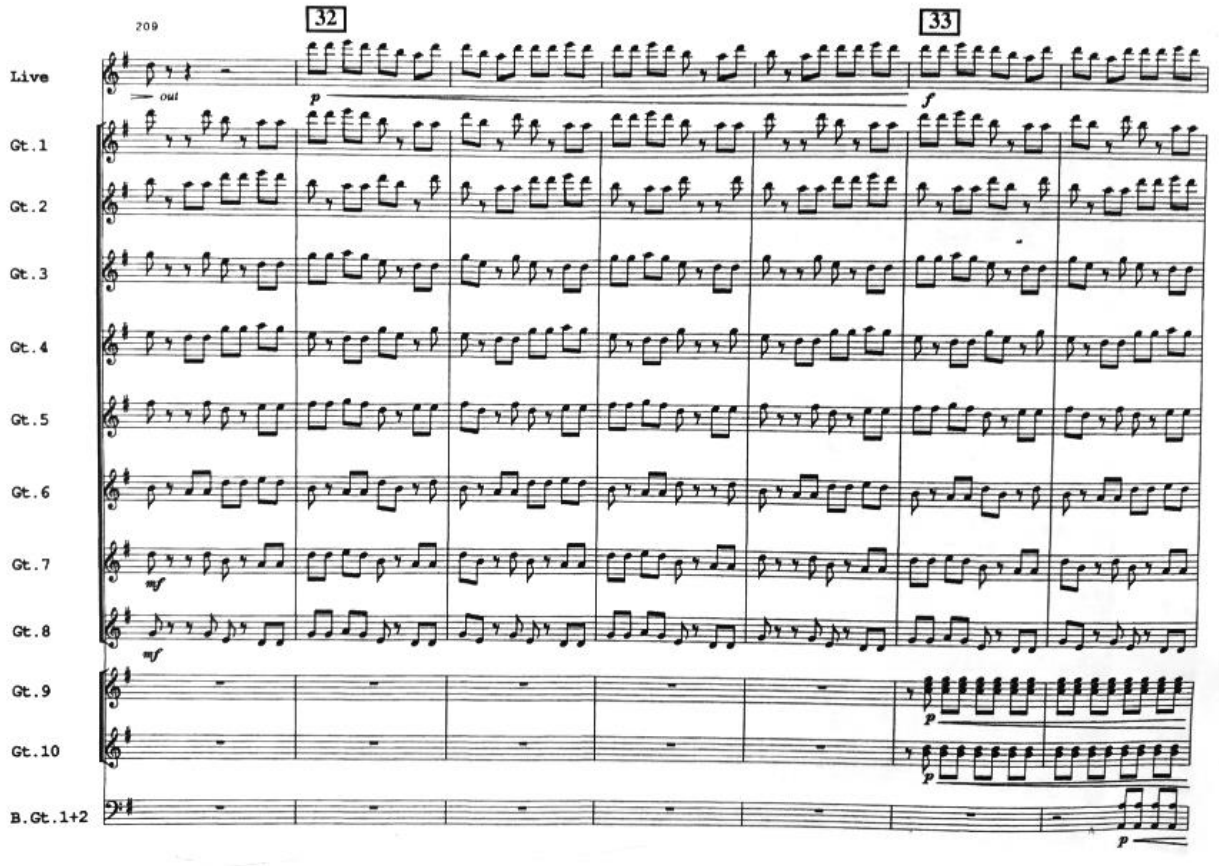

Figure 23. Electric Counterpoint, Movement I, mm. 209–215.

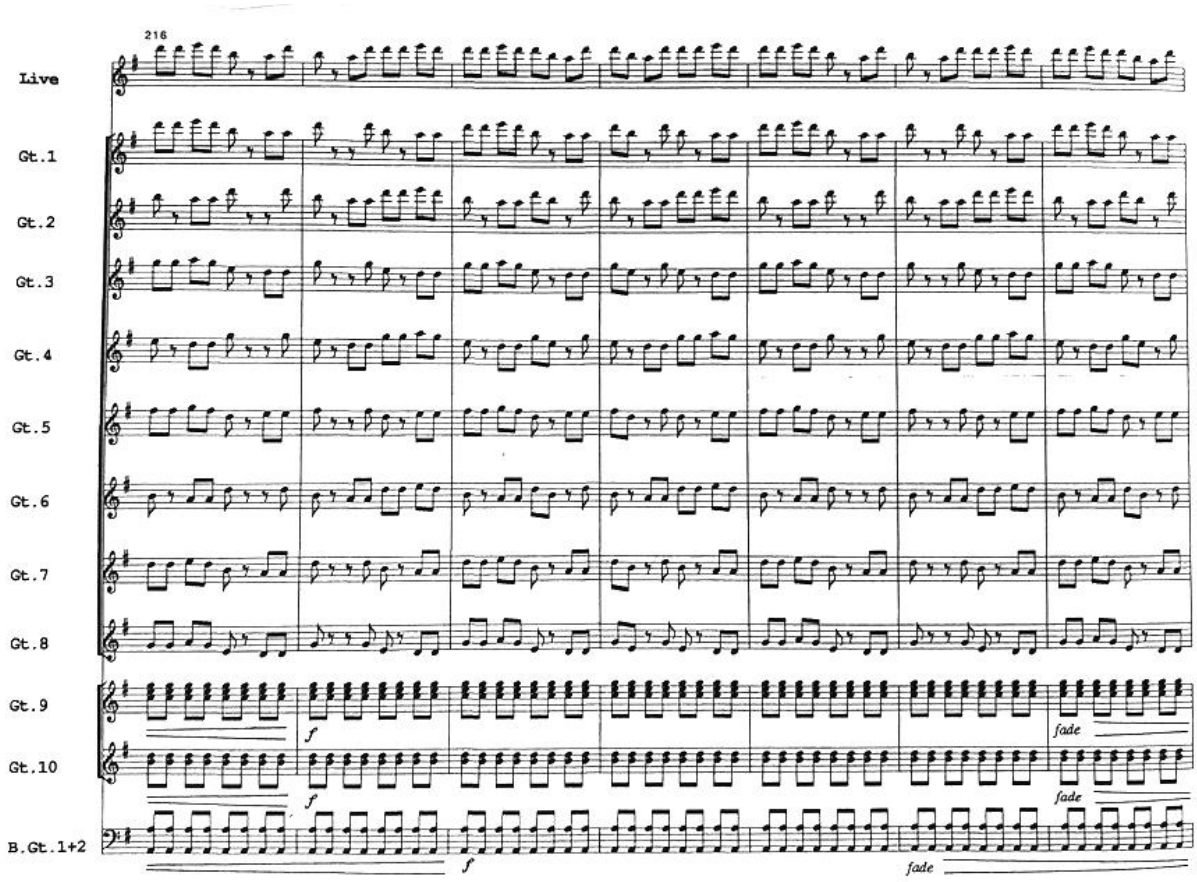

Figure 24. Electric Counterpoint, Movement I, mm. 216–222.

Hansen writes how electronic (specifically, digital) media create a mixed reality for the participant-spectator, with different arenas for action that are accessed by an embodied participant. But here, this alternate space, the space beyond the hocketing, is *not* accessible. We know it's out there: we can hear the reverberation that signals its extent, but we can't reach it. This may seem an odd statement to make about a music listening experience, but consider the following question: in terms of movement, does the dense counterpoint afford dancing in place, or prancing across wide spaces like a gymnast performing a floor routine? I believe most listeners would say the former. This is not a claim that there are movements we make in response to music that are unsullied by cultural conventions and social norms.

But as demonstrated by Kozak, our affective responses to music as reflected in bodily expression are not always arbitrary ones predetermined through enculturation and socialization. Consequently, our bodily responses can be drawn upon in support of a particular hearing of a piece.[245] I maintain that the movements solicited by this section of *Electric Counterpoint* are ones that are restricted to peripersonal space and do not involve locomotion, and that this, in turn, is indicative of the structuring of space accomplished by the music.

6.6 Space and Bass and Mixed Reality

This moment of high activity and spatial constraint in measures 210–213 is an appropriate point for a dénouement. That dénouement is delivered with the onset of the third and final section of the movement, which brings the bass guitars back into the texture along with the strummed chords in guitar parts 9 and 10 at measure 214 (Figure 22; 4:20). The addition of these two harmonic parts, along with the electric basses playing in unison, creates the effect of a breaking through and breaking free of the screen erected by guitars 1 though 8 in conjunction with the soloist. This sensation is achieved in part by the way in which the new elements partially mask the hocketing pattern. The added instrumental strata make it difficult to hear the parts featured in the second section as a patterned construction, made up of strategically overlaid melodic loops. The reflected waveforms are masked and the prominence of the upper parts wanes. In addition, these new parts crescendo to *forte*, against the *mezzo-forte* of the others. The solo guitar line is also playing

245 Mariusz Kozak, "Listeners' Bodies in Music Analysis: Gestures, Motor Intentionality, and Models," *Music Theory Online* 21, no. 3 (2015), http://www.mtosmt.org/issues/mto.15.21.3/mto.15.21.3.kozak.html.

forte here, and for the first time a homophonic texture arises that consists of a single melody with harmonic accompaniment. But more importantly for the purposes of the release of tension is the introduction of the low octaves played by the electric bass parts. Taken together, these elements effect the sonic "breakthroughs" that rise up throughout the movement's closing section.[246]

When sound is felt in addition to heard, the effect on the listener is a powerful one. The surrounding mesh of sound that this piece creates in higher frequencies is supplemented by vibrations transmitted to the body by any surface the listener is in contact with, such as a chair, a table, or the floor. This haptic variety of multi-modal music listening is one that we are so used to that it very often goes by unremarked in music theory. Much of the time, a newly introduced bass part is heard simply as additional loudness. Yet the sound of the bass is enveloping in a way that, for instance, guitar parts 1 through 8 are not. While the spatial parameters established at the beginning of the piece have remained constant, the available space for movement has been modified—we have been provided access to the full extent of the space indicated by the sound. In terms of the "I can" of Merleau-Ponty, we now have more space available to us to explore.[247] The haptic component contributes to the spatiality of our situation by indexing sound's materiality and our own corporeality. Our body is enlisted in the listening event and forecloses any pretense of a purely intellectual appreciation of the music. In contrast to "Kailasha," the

246 As Leydon (2012) emphasizes, different performance choices highlight different possibilities for listener engagement. Rather than releasing or breaking through the surrounding hocketing, Sean Shibe crafts his performance/recording of *Electric Counterpoint* in such a way that the bass adds yet another layer of complexity to the hocketing sounds. Rather than release, we are further enmeshed. Like bass sounds in general, these voices are more difficult to localize, but this position-less aspect is not exploited for its spatial qualities as it is in the Metheny recording. Sean Shibe, *softLOUD: Music for Acoustic & Electric Guitars*, Delphian (DCD34213).

247 Merleau-Ponty, *Phenomenology of Perception*, 2012, 139.

music reaches out to touch us, rather than the other way around. In this movement's preceding section, the reverberation (whether added electronically or not) alerted us to a larger space out of our reach, beyond the hocketing. This was a space that we determined earlier was not in use. However, if the hocketing texture has been dissolved, if the surfaces of our listening space are vibrating, and if we proceed from feeling confined by musical sound to feeling enveloped in it, then we can confidently assert that the music has succeeded in structuring the space around us.

Hansen notes how the introduction of the tactile elements into visual media serves as a supplement to the image, independent of its referent and capable of conferring "a more concrete sense of 'reality' on it." It "catalyzes" the creation of "a singular affective experience."[248] In such cases, the body is engaged in a process of unifying the elements in its purview, or in Merleau-Pontian terms, of establishing equilibrium. The experience is singular, rather than fragmented, because of the body's unique and persistent role as the constant, fixed background that frames our experience and allows the things we perceive to show up.[249] Our body represents an irreducible ground from which we approach the world.

This merging of environments is another example of the mixed reality that Hansen theorizes and which I explored in Part I with respect to the electronic jazz tracks of Miles Davis and his studio engineer Teo Macero.[250] Because the bass is so pervasive throughout the environment (as opposed to, say, virtual or augmented reality apps that require us to

248 Hansen claims this catalyzation of experience extends beyond the tactile supplement: touch, feeling, affect, movement, and perception all operate in concert to reveal the aesthetic image. Mark B. N. Hansen, *New Philosophy for New Media* (Cambridge, Mass: MIT Press, 2004), 39.

249 "The body forms an ultimate background, an absolute here, in relation to which all perceptual experience must be oriented." Hansen, *Bodies in Code*, 2005, 5.

250 This is a variety of what Hans-Georg Gadamer calls a "fusion of horizons." Hans-Georg Gadamer, *Truth and Method* (New York: Continuum, 1994), 306–7.

look directly at a screen, or virtual reality goggles, that block out the everyday visual environment), and because it shakes our bodies and the chair we sit in (or floor we stand on, or couch we lie on), musical space fuses with our own, an integration accomplished via the body that is the hallmark of Hansen's model. However, in listening to recorded music, the sonic realities we have access to are not just mixed, but merged. We don't hear Radigue's music as separate from actual room sound. Similarly, we don't hear the sound of a tam-tam as a simulation, but as suffused in our own space. Neither is the spatial structuring accomplished by *Electric Counterpoint* an imaginary simulation. It is a part of our inhabited reality.

6.7 Motor-Intentionality and the Generation of Space

Clearly, the feeling of breaking free, or of breaking through, that can be experienced in the final section of the first movement is not only a sensation but also a meaning, a meaning that is inaccessible through recourse to the score alone. The meaning we find here also has a narrative aspect, as we cannot have an experience of breaking out unless there has been a previous hemming in. The two events are sequenced and thus form a miniature plot. But this world is disclosed to us differently than those which are revealed through literature. As we noticed with Radigue's "Kailasha," sounds give space a texture we can grasp with respect to our own proprioceptive awareness. While *Electric Counterpoint* does not afford the same opportunities for detecting non-linear variations in our surrounding space, we take in its musical components in the same way. It is with respect to our habituated understanding of our own bodies in space that we hear sounds as near or far.

And here, it is the potential for movement indicated by the music that gives us this spatial awareness. We know from the reverberations that there is space beyond the peripersonal even as we contend with those sounds close at hand.

For Merleau-Ponty, perception involves "motor-intentionality," claiming that our motions are always directed toward something, even if inchoately. This is the crux of his concept of embodiment: not only is thought intentional (as in Husserl's phenomenology), so is perception and movement. Scott Marratto reminds us that for Merleau-Ponty, motor-intentionality is the awareness of the body and its abilities expressed by the phrase "I can." It is an intentionality that serves as the underlying ground to our perceptions, a ground we always possess and that always imposes its own schema onto our surroundings.

> The phenomenon of space, and, in particular, depth, must be understood in terms of the bodily 'I can,' as the arena of an organism's possibilities for acting: the space of affordances.... Space as it is rooted in the spatiality of one's own situated body.[251]

This "I can" serves as a set of norms to organize our perceptions. These norms impose their own Gestalt structure upon our world—one that, according to Merleau-Ponty, is primordial. As Taylor Carman explains:

> What makes motor "intentionality" worthy of the name is precisely its normativity, the felt rightness and wrongness of the various bodily attitudes we unthinkingly assume and maintain throughout our waking (and sleeping) lives. Felt differences between manifestly better and worse attitudes mark differences between right and wrong, or true and false, perceptual appearances: the words on the chalkboard are indistinct, so I squint and crane my neck to see them better; the voice at the back of the room is muffled, so I lean forward and put my hand to my ear; the sweater looks brown until I hold it directly under the light and see that it is green.[252]

251 Marratto, *The Intercorporeal Self*, 43.
252 Carman, *Merleau-Ponty*, 110.

These felt differences in the relationship between body and world mean that we do not hear spatial configurations in an imaginary or metaphorical or analogical fashion, but in terms of the concrete opportunities they allow us for engaging with the things we encounter within it. Space appears to us through our own affective responses to the potential of bodily movement: that is, through motor intentionality.

The things we encounter force us to reckon with our surroundings because they demand we take up a perspective if we are to approach them at all. That activity of engagement—not simply the perspective-taking itself but also the *negotiation* of a perspective, the grasping for optimal grip—takes us out of the habitual and opens us to the world in new ways, for in order to engage with music we must grapple with it. As Merleau-Ponty writes:

> "It is the price we must pay to have a literature, that is a conquering language which introduces us to *unfamiliar* perspectives instead of confirming us in our own…. It is essential to what is true to be presented first and forever in a movement which throws our image of the world *out of focus*."[253]

If we can manage to truly detach from the Cartesian perspective of a sovereign and self-sufficient mind receiving and assembling incoming sensations, then we can understand logically that sound and music instantiate a world, a world that is coextensive with whatever situation we find ourselves in. It is through engaging with this proposed world that we are invited to "take up the gesture" and make it our own. This is true whether we

253 Maurice Merleau-Ponty, "Indirect Language and the Voices of Silence," in *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, ed. Galen A. Johnson, trans. Michael B. Smith, Northwestern University Studies in Phenomenology and Existential Philosophy (Evanston, IL: Northwestern University Press, 1993), 113–14. Italics in original.

are engrossed in the type of studied contemplation called “structural listening,” or decide to join in the musicking by dancing, marching, or singing along: we are always active participants. Through employing a phenomenological approach to perception and meaning that takes the body into account, we can better understand how music-theoretical structures such as texture and timbre construct the affective experiences that we value in musical works.

References

Ashby, Arved Mark. *Absolute Music, Mechanical Reproduction*. Berkeley: University of California Press, 2010.

Bachelard, Gaston. *The Poetics of Space*. New York: Orion Press, 1964.

Bakhtin, Mikhail. *Toward a Philosophy of the Act*. Austin: University of Texas Press, 1993.

Barnett, Kyle S. "Furniture Music: The Phonograph as Furniture, 1900–1930." *Journal of Popular Music Studies* 18, no. 3 (December 2006): 301–24. https://doi.org/10.1111/j.1533-1598.2006.00096.x.

Barthes, Roland. *Mythologies*. Translated by Annette Lavers. New York: Noonday Press, 1972.

Berry, Wallace. *Structural Functions in Music*. New York: Dover, 1987.

Besseler, Heinrich. "Appendix: Fundamental Issues of Musical Listening (1925)." Translated by Matthew Pritchard and Irene Auerbach. *Twentieth Century Music* 8, no. 1 (March 2011): 49–70.

Blauert, Jens. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, Mass: MIT Press, 1997.

Böhme, Gernot. "Atmosphere as the Fundamental Concept of a New Aesthetics." *Thesis Eleven* 36, no. 1 (August 1, 1993): 113–26. https://doi.org/10.1177/072551369303600107.

———. *Atmospheric Architectures: The Aesthetics of Felt Spaces*. Translated by Tina Engels-Schwarzpaul. New York, NY: Bloomsbury Academic, 2017.

"Brant, Henry | Grove Music." Accessed April 11, 2018. http://www.oxfordmusiconline.com/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000003850.

Brinkema, Eugenie. *The Forms of the Affects*. Durham: Duke University Press, 2014.

Bryant, David. "The 'Cori Spezzati' of St Mark's: Myth and Reality." *Early Music History* 1 (1981): 165–86.

Carman, Taylor. *Merleau-Ponty*. New York: Routledge, 2008.

Casey, Edward S. “How to Get from Space to Place in a Fairly Short Stretch of Time: Phenomenological Prolegomena.” In *Senses of Place*, 13–52. Santa Fe, N.M.: School of American Research Press, 1996.

———. *The Fate of Place: A Philosophical History*. Berkeley: University of California Press, 1997.

———. *The Fate of Place : A Philosophical History*. Berkeley: University of California Press, 1997.

Chabot, Pascal. *The Philosophy of Simondon: Between Technology and Individuation*. London: Bloomsbury Academic, 2013.

Clarke, Eric F. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. New York: Oxford University Press, 2005.

Dax, Max. “Eliane Radigue: An Interview.” *Telekom Electronic Beats* (blog), October 11, 2012. https://www.electronicbeats.net/eliane-radigue-an-interview/.

DeGraaf, Leonard. “Confronting the Mass Market: Thomas Edison and the Entertainment Phonograph.” *Business and Economic History* 24, no. 1 (1995): 88–96.

DeNora, Tia. *Music in Everyday Life*. New York: Cambridge University Press, 2000.

Dockwray, Ruth, and Allan F. Moore. “Configuring the Sound-Box 1965–1972.” *Popular Music* 29, no. 2 (2010): 181–97.

———. “Configuring the Sound-Box 1965–1972.” *Popular Music* 29, no. 2 (2010): 181–97.

Doyle, Peter. *Echo and Reverb: Fabricating Space in Popular Music, 1900-1960*. Middletown, Conn: Wesleyan University Press, 2005.

———. *Echo and Reverb: Fabricating Space in Popular Music, 1900-1960*. Middletown, Conn: Wesleyan University Press, 2005.

Dunsby, Jonathan. “Considerations of Texture.” *Music & Letters* 70, no. 1 (1989): 46–57.

Dyson, Frances. *Sounding New Media: Immersion and Embodiment in the Arts and Culture*. Berkeley: University of California Press, 2009.

Edison, Thomas A. “The Phonograph and Its Future.” *The North American Review* 126, no. 262 (1878): 527–36.

Eidsheim, Nina Sun. *Sensing Sound: Singing & Listening as Vibrational Practice*. Durham: Duke University Press, 2015.

"Electric Guitar | Grove Music." Accessed May 24, 2018. http://www.oxfordmusiconline.com/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-1002256412.

"Electric Guitar | Grove Music." Accessed May 24, 2018. http://www.oxfordmusiconline.com/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000008690.

Emmerson, Simon. *Living Electronic Music*. Burlington, VT: Ashgate, c2007.

Fahy, Frank. *Foundations of Engineering Acoustics*. Oxford: Elsevier Science & Technology, 2000.

Feld, Steven. "They Repeatedly Lick Their Own Things." *Critical Inquiry* 24, no. 2 (1998): 445–72.

Fessel, Pablo. "Hacia Una Caracterización Formal Del Concepto de Textura." *Revista Del Instituto Superior de Música*, no. 5 (1996): 75–93.

Fletcher, N. H. "The Nonlinear Physics of Musical Instruments." *Reports on Progress in Physics* 62, no. 5 (January 1999): 723–764.

Freeman, Phil. *Running the Voodoo Down: The Electric Music of Miles Davis*. Berkeley, CA: Publishers Group West, 2005.

Frith, Simon. *Performing Rites: On the Value of Popular Music*. New York: Oxford University Press, 1996.

Gadamer, Hans-Georg. *Philosophical Hermeneutics*. Berkeley: University of California Press, 2004.

———. *Truth and Method*. New York: Continuum, 1994.

Gibson, James J. *The Ecological Approach to Visual Perception*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1986.

Guck, Marion A. "Music Loving, or the Relationship with the Piece." *The Journal of Musicology* 15, no. 3 (1997): 343–52.

Hansen, Mark B. N. *Bodies in Code: Interfaces with Digital Media*. London: Routledge, 2005.

———. *Bodies in Code: Interfaces with Digital Media*. London: Routledge, 2005.

———. *New Philosophy for New Media*. Cambridge, Mass: MIT Press, 2004.

———. *New Philosophy for New Media*. Cambridge, Mass: MIT Press, 2004.

Hatten, Robert S. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington, IN: Indiana University Press, 2004.

Iverson, Jennifer. "Invisible Collaboration: The Dawn and Evolution of Elektronische Musik." *Music Theory Spectrum* 39, no. 2 (2017): 200–22.

Johnson, Mark. *The Meaning of the Body: Aesthetics of Human Understanding*. Chicago: University of Chicago Press, 2007.

Johnston, Blair. "Between Romanticism and Modernism and Postmodernism: George Crumb's Black Angels." *Music Theory Online* 18, no. 2 (June 1, 2012). http://mtosmt.org/issues/mto.12.18.2/mto.12.18.2.johnston.html.

Kane, Brian. *Sound Unseen: Acousmatic Sound in Theory and Practice*. New York, NY: Oxford University Press, 2014.

Katz, Mark. *Capturing Sound: How Technology Has Changed Music*. Berkeley: University of California Press, 2010.

Keightley, Keir. "'Turn It Down!' She Shrieked: Gender, Domestic Space, and High Fidelity, 1948-59." *Popular Music* 15, no. 2 (1996): 149–77.

Kelly, Sean Dorrance. "Seeing Things in Merleau-Ponty." In *The Cambridge Companion to Merleau-Ponty*, edited by Taylor Carman and Mark B.N. Hansen. New York: Cambridge University Press, 2005.

Klett, Joseph. "Sound on Sound: Situating Interaction in Sonic Object Settings." *Sociological Theory* 32, no. 2 (2014): 147–61.

Kozak, Mariusz. "Listeners' Bodies in Music Analysis: Gestures, Motor Intentionality, and Models." *Music Theory Online* 21, no. 3 (2015). http://www.mtosmt.org/issues/mto.15.21.3/mto.15.21.3.kozak.html.

Lacasse, Serge. "Persona, Emotions and Technology: The Phonographic Staging of the Popular Music Voice." *CHARM Symposium 2: Towards a Musicology of Production*, 2005, 11.

———. “Voice and Sound Processing: Examples of Mise En Scène of Voice in Recorded Rock Music.” *Popular Musicology Online*, no. 5 (2000). http://www.popular-musicology-online.com/issues/05/lacasse.html.

Lefebvre, Henri. *The Production of Space*. Cambridge, Mass., USA: Blackwell, 1991.

Leydon, Rebecca. “Clean as a Whistle: Timbral Trajectories and the Modern Musical Sublime.” *Music Theory Online* 18, no. 2 (June 2012). https://mtosmt.org/issues/mto.12.18.2/mto.12.18.2.leydon.html.

Margulis, Elizabeth. *On Repeat: How Music Plays the Mind*. New York, NY: Oxford University Press, 2014.

Marratto, Scott L. *The Intercorporeal Self: Merleau-Ponty on Subjectivity*. Albany: State University of New York Press, 2012.

Massumi, Brian. *Parables for the Virtual : Movement, Affect, Sensation*. Durham [N.C.]: Duke University Press, 2002.

Merleau-Ponty, Maurice. “Eye and Mind.” In *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, edited by Galen Johnson, translated by Michael Smith, 121–49. Evanston, IL: Northwestern University Press, 1993.

———. “Indirect Language and the Voices of Silence.” In *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, edited by Galen Johnson, translated by Michael Smith, 76–120. Evanston, IL: Northwestern University Press, 1993.

———. “Indirect Language and the Voices of Silence.” In *The Merleau-Ponty Aesthetics Reader: Philosophy and Painting*, edited by Galen A. Johnson, translated by Michael B. Smith, 76–120. Northwestern University Studies in Phenomenology and Existential Philosophy. Evanston, IL: Northwestern University Press, 1993.

———. *Phénoménologie de La Perception*. Paris: Gallimard, 1945.

———. *Phenomenology of Perception*. Translated by Colin Smith. New York: Routledge, 2002.

———. *Phenomenology of Perception*. Translated by Donald A. Landes. New York: Routledge, 2012.

———. *Phenomenology of Perception*. Translated by Donald A. Landes. New York: Routledge, 2012.

Metzger, Wolfgang. “Optische Untersuchungen am Ganzfeld. II. Mitteilung: Zur Phänomenologie des homogenen Ganzfelds [Optical Investigations of the Ganzfeld. II. Toward the Phenomenology of the Homogeneous Ganzfeld].” *Psychologische Forschung* 13, no. 1 (January 1, 1930): 6–29. https://doi.org/10.1007/BF00406757.

Meyer, Leonard B. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956.

Michels, Christoph. “Researching Affective Atmospheres.” *Geographica Helvetica* 70, no. 4 (October 5, 2015): 255–63. https://doi.org/10.5194/gh-70-255-2015.

Miller, Paul. “An Adventure into Outer Space: Stockhausen’s Lichter—Wasser and the Analysis of Spatialized Music.” *Perspectives of New Music* 50, no. 1–2 (2012): 342–92. https://doi.org/10.7757/persnewmusi.50.1-2.0342.

Milner, Greg. *Perfecting Sound Forever: An Aural History of Recorded Music*. New York: Faber and Faber, 2009.

Morris, David. *The Sense of Space*. Albany: State University of New York Press, 2004.

———. *The Sense of Space*. Albany: State University of New York Press, 2004.

Moylan, William. “An Analytical System for Electronic Music.” D.A., Ball State University, 1983. https://search-proquest-com.ezproxy.cul.columbia.edu/pqdtglobal/docview/303122930/AF6A649A6D2F4856PQ/1.

———. “An Analytical System for Electronic Music.” D.A., Ball State University, 1983. http://search.proquest.com/pqdtglobal/docview/303122930/abstract/A316BAFA0E854DA0PQ/5.

———. “The Elements of Sound and Audio Recording.” In *Understanding and Crafting the Mix: The Art of Recording*. Boston: Elsevier/Focal Press, 2007.

———. *Understanding and Crafting the Mix [Electronic Resource]: The Art of Recording*. Boston: Elsevier/Focal Press, 2007.

———. *Understanding and Crafting the Mix: The Art of Recording*. Boston: Elsevier/Focal Press, 2007.

Nudds, Matthew. “Sounds and Space.” In *Sounds and Perception: New Philosophical Essays*, edited by Matthew Nudds and Casey O’Callaghan. New York: Oxford University Press, 2009.

Ockeghem, Johannes. *Missa Mi-Mi*. Edited by Heinrich Besseler. Wolfenbüttel: Möseler Verlag, n.d.

Ricoeur, Paul. "Metaphor and the Main Problem of Hermeneutics." *New Literary History* 6, no. 1 (1974): 95–110. https://doi.org/10.2307/468343.

Romdenh-Romluc, Komarine. *Routledge Philosophy Guidebook to Merleau-Ponty and Phenomenology of Perception*. New York: Routledge, 2011.

Rumsey, Francis. "Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm." *Journal of the Audio Engineering Society* 50, no. 9 (September 15, 2002): 651–66.

Schrimshaw, Will. *Immanence and Immersion: On the Acoustic Condition in Contemporary Art*. New York: Bloomsbury Academic, 2017.

Schütze, Paul. "Surround Sound." *Frieze*, 2011. https://frieze.com/article/surround-sound.

Smart, James R. "Emile Berliner and Nineteenth-Century Disc Recordings." *The Quarterly Journal of the Library of Congress* 37, no. 3/4 (1980): 422–40.

Steege, Benjamin. *Helmholtz and the Modern Listener*. New York: Cambridge University Press, 2012.

Steinby, Liisa. "Hermann Cohen and Bakhtin's Early Aesthetics." *Studies in East European Thought* 63, no. 3 (2011): 227–49.

Sterne, Jonathan. *The Audible Past: Cultural Origins of Sound Reproduction*. Durham: Duke University Press, 2003.

Stockhausen, Karlheinz, and Jerome Kohl. "Electroacoustic Performance Practice." *Perspectives of New Music* 34, no. 1 (1996): 74–105. https://doi.org/10.2307/833486.

"Texture | Grove Music." Accessed June 18, 2019. https://www-oxfordmusiconline-com.ezproxy.cul.columbia.edu/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000027758.

Valiquet, Patrick. "The Spatialisation of Stereophony: Taking Positions in Post-War Electroacoustic Music." *International Review of the Aesthetics and Sociology of Music* 43, no. 2 (2012): 403–21.

Xie, Bosun. *Head-Related Transfer Function and Virtual Auditory Display*. Plantation, Florida: J. Ross Publishing, 2013.

Zagorski-Thomas, Simon. “The Stadium in Your Bedroom: Functional Staging, Authenticity and the Audience-Led Aesthetic in Record Production.” *Popular Music* 29, no. 2 (2010): 251–66.

Zayaruznaya, Anna. “Intelligibility Redux: Motets and the Modern Medieval Sound.” *Music Theory Online* 23, no. 2 (June 1, 2017). http://mtosmt.org/issues/mto.17.23.2/mto.17.23.2.zayaruznaya.html.

Zvonar, Richard. “A History Of Spatial Music.” *EContact* 7, no. 4 (2004). https://econtact.ca/7_4/zvonar_spatialmusic.html.