

Copyright
by
Ian Wyatt Gerg
2015

**The Dissertation Committee for Ian Wyatt Gerg certifies that this is the approved
version of the following dissertation:**

**The Virtual Observing Agent in Music:
A Theory of Agential Perspective as Implied by Indexical Gesture**

Committee:

Robert S. Hatten, Supervisor

Byron Almén

Eric Drott

Edward Pearsall

Katrin Erk

**The Virtual Observing Agent in Music:
A Theory of Agential Perspective as Implied by Indexical Gesture**

by

Ian Wyatt Gerg, B.S., M.Music

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2015

Dedication

For my parents

Acknowledgements

I would like express my gratitude to the many teachers, friends, and family members whose mentorship and support have enabled the completion of this dissertation.

First and foremost, I must thank my advisor, Robert S. Hatten, for his guidance, encouragement, and patience during our weekly meetings over the past number of years. The value of his insights is only rivaled by the depth of his kindness. The music theory and musicology faculties should also know my appreciation for my time with them. Their effectiveness as pedagogues and mentors is something to be emulated. I owe gratitude to my fellow graduate students who have been unselfish with their encouragement and friendship. I am further thankful for the generosity of the Kent Kennan Endowed Graduate Fellowship in Music Theory (2008-2009; 2012-2013).

A number of others played a significant role in this project long before its inception. My earliest music teachers, Ray Knight and Joe Sensor, let me know the value of personal expression. Mary Lou Cartwright and Mary Steele taught me the art of language. I owe a debt of gratitude to Mark Lusk and Eric McKee for opening my mind as a musician, a scholar, and a person. I am also appreciative of the inspiration and encouragement from Jeff Vollmer, who first showed me a different way to approach music and gave me something to reach for as an undergraduate.

The members of my family, particularly Lois and Dick Sadley, have provided me with a tremendous amount of support. Their interest in my work has made it easy to walk this uncommon path. I must also express my appreciation for my brother, Isaac. Through his example and mentorship, I learned the value of curiosity and diligence at an early age.

Finally, and above all, I am grateful for the love and guidance selflessly offered by my parents. They made this project possible each and every day. This is for them.

**The Virtual Observing Agent in Music:
A Theory of Agential Perspective as Implied by Indexical Gesture**

Ian Wyatt Gerg, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Robert S. Hatten

The human body is inseparable from our understanding of music. Through embodied cognition, listeners conceptualize music as performed action. We find evidence of this in our most fundamental musical language. “High” pitches resonate high in a singer’s head, while “fast” rhythms resemble fast bodily movement. Scholars have followed the entailments of these metaphors in recent decades, developing theories of bodily gesture (Hatten 2004, Lidov 2005) and physical mimesis (Cox 2011). These hold that the bodily movement that we hear in music can imitate the physical gestures that we use in everyday communication (e.g., waving, nodding, bowing, or sighing). This has its own entailments; most fundamentally, it implies the presence of a virtual, human-like agent within music that is similar to the “virtual persona” theorized by Edward T. Cone (1974). In other words, in perceiving musical sounds as imitative of physical movement and gesture, we infer the presence of a virtual agent who enacts them.

This dissertation extends these theories, demonstrating that musical gestures can be mimetic of *indexical* somatic movements—that is, bodily movements of pointing,

looking, striving, and reaching. These indexical gestures suggest the presence of a virtual *observing* agent. The virtual observing agent acts a lens through which we, the listener, can experience the interior world (diegesis) of a work. This leads us to embody a single and more individualized perspective on the musical representation. I explore the implications of indexical gesture and perspective with an examination of music from the common practice period. Moreover, I bring the theory of virtual observing agency together with theories of musical narrative and emotion.

Table of Contents

List of Examples	x
List of Figures	xi
Chapter 1 Toward a Theory of the Virtual Observing Agent	1
An Introduction to Gesture	1
From Physical Gesture to Musical Gesture	2
Scholarly Precedents for a Theory of Gesture	5
The Present State of Gesture Theory	13
How Agency Is Implied	17
The Implication of the Diegesis	19
A New Theory of Gesture and Agential Perspective	22
Indexical Gestures	27
Breathing or Gazing: Lobbying for the Indexical Gesture	41
Conclusion	43
Chapter 2 Mathematical and Logical Formulation of Indexical Gestures	45
Introduction	45
Part I Evidence for the Gazing Gesture	46
Part II Inferring a Virtual Observing Agent	59
Markedness	60
From Listener to Virtual Agent	65
Case Study 1: Chopin, Prelude in B Minor, Op. 26, no. 6	65
Case Study 2: Chopin, Barcarolle in F-sharp Major, Op. 60	80
Conclusion	89
Chapter 3 Other Observing Gestures and Perspectives	90
Introduction	90
Part I Yearning as Indexical Gesture	91
Indexical Gesture in Beethoven, Piano Concerto No. 5, ii	91
Teleological Melodic Ascent in Indexical Gesture	96
Yearning	109

Yearning as Performance Gesture	110
Yearning and Yielding.....	112
Yielding as Performance Gesture	115
Part II Observing from Multiple Perspectives	121
One Object, Multiple Perspectives	122
Multiple Perspectives as Expressive Doubling.....	127
Multiple Perspectives in Debussy, <i>La cathédrale engloutie</i>	130
Conclusion	140
Chapter 4 Virtual Observing Agency, Emotion, and Narrative.....	141
Introduction.....	141
From Motion to Emotion	141
Music and Emotion: A Physical Experience	145
Music and Emotion: A Cognitive Experience	149
Modes of Emotional Engagement.....	152
Musical Frames and Narrative	161
The Narrator in Mendelssohn, <i>Song without Words</i> , Op. 30, no. 3	165
The Narrator in Dukas, <i>The Sorcerer's Apprentice</i>	172
Observation and Emotion in Berlioz, <i>Harold in Italy</i> , ii	179
Conclusion	186
Epilogue	187
Bibliography	193
Vita.....	200

List of Examples

Example 1.1:	Chopin, Prelude in B Minor, Op. 28, no. 6, mm. 1-2	26
Example 1.2:	Beethoven, Symphony No. 2, Op. 36, iv, mm. 1-2.....	32
Example 1.3	Liszt, <i>La lugubre gondola</i> , S. 200/1, accompaniment	35
Example 1.4:	Liszt, <i>La lugubre gondola</i> , S. 200/1, mm. 94-120.....	37
Example 1.5:	Schumann, Symphony No. 3 in E-flat Major, iv, mm. 1-5.....	39
Example 2.1:	J.S. Bach, English Suite No. 1, Sarabande, mm. 1-4.....	72
Example 2.2	Saint-Saëns, <i>Samson and Delilah</i> , Bacchanale, recitative.....	79
Example 2.3:	Chopin, Barcarolle in F-sharp Major, Op. 60, mm. 1-6.....	81
Example 3.1:	Beethoven, <i>Emperor Concerto</i> , Op. 73, ii, mm. 1-17.....	92
Example 3.2:	Kent and Burton, “The White Cliffs of Dover”	99
Example 3.3:	Beethoven, Piano Sonata in A Major, Op. 101, i.....	106-08
Example 3.4:	Beethoven, Piano Sonata in C Minor, Op. 13, i, m. 1.....	118
Example 3.5:	Bacon, <i>Three Studies of Lucian Freud</i>	124
Example 3.6:	Beethoven, Piano Sonata in F-sharp Major, Op. 78, i and ii	128
Example 3.7:	Debussy, <i>La cathédrale engloutie</i> , Préludes Bk. I, no. 10..	131-35
Example 4.1:	Haydn, <i>Surprise Symphony</i> , ii, mm. 1-16	147
Example 4.2:	Mozart, Requiem, “Lacrimosa,” mm. 1-2.....	149
Example 4.3:	Mendelssohn, <i>Song without Words</i> , Op. 30, no. 3.....	166
Example 4.4:	Berlioz, <i>Harold in Italy</i> , i, mm. 38-45, solo viola	180
Example 4.5:	Berlioz, <i>Harold in Italy</i> , ii, mm. 54-69	183

List of Figures

Figure 1.1:	Meyer's equation for determining conformance	10
Figure 2.1:	Parabola.....	47
Figure 2.2:	Inverted Parabola	48
Figure 2.3:	Path of motion of object.....	50
Figure 2.4:	Right triangle placed against path of motion of object	53
Figure 2.5:	Path of motion shifted 10 feet right on the x-axis.....	54
Figure 2.6:	Angle of vision at 10 feet from point of propulsion	56
Figure 2.7:	Rate of movement in observer's viewing gesture.....	58
Figure 2.8:	Kopp's description of Chopin, Barcarolle, mm. 1-3	83
Figure 3.1:	Neumeier's middleground of Beethoven, Op. 101, i.....	102
Figure 3.2:	Reproduction of Zuckerkandl's melodic tendency diagram	103
Figure 3.3:	Formal layout of <i>La cathédrale engloutie</i> (Ericourt).....	137
Figure 4.1:	Transmission of Emotion: a) Empathy and b) Sympathy	156
Figure 4.2:	Flow of diegetic information to the listener.....	182

CHAPTER 1

TOWARD A THEORY OF THE VIRTUAL OBSERVING AGENT

An Introduction to Gesture

The opening motive of Beethoven's Symphony No. 5 in C minor is perhaps the best-known musical gesture in all of Western classical music. The urgency of the rhythm and the intensity of its sequential repetition that is set in the minor mode depict both an unprecedented potency and dire inevitability. In *Beethoven as I Knew Him*, Anton Schindler recalls:

The composer himself provided the key to these depths when one day, in this author's presence, he pointed to the beginning of the first movement and expressed in these words the fundamental idea of his work: 'Thus Fate knocks at the door!' (1966 [1840, 1860]: 147)

Of course, given Schindler's known mendacious tendencies, this anecdote is likely a romantic fabrication. Regardless, it takes no interpretive leap to imagine the music representing what Schindler describes. We hear the first four pitches of the motive sound as a forceful hand rapping against a large door, while its repeat at a lower pitch level suggests its urgency. After a brief pause, the knocking proceeds with reckless abandon as the impetuosity of Fate develops throughout the movement.

In hearing the motive as knocking, we are describing musical sounds as though they comprise a physical gesture—a real bodily movement. This leads to two important implications. First, the knocking gesture implies that the action is that of a human *agent*, in this case an anthropomorphized Fate; our understanding of a physical action assumes the existence of an entity that performs this action. However, because the gesture and the agent who performs it only exist in our conceptualization of the music, not in reality, we additionally infer the existence of a *diegesis*—a virtual world distinct from the real, material space in which we, as listeners, exist. Although we might take these three inferences for granted, their entailments prove to be valuable to interpretation.

Before proper definitions of a musical gesture, an agent, and the diegesis can be offered, it is necessary first to understand the components of a physical gesture since the latter is the source of the musical metaphor.

From Physical Gesture to Musical Gesture

As humans we engage in physical somatic gestures daily because they exist as significant components of communication. For example, when greeting someone whom we have not recently seen, we extend a hand to invite a handshake. While in traffic, we may wave with gratitude to a driver who has let us proceed first at an intersection. Other human gestures are less voluntary, such as when we yelp in pain after stubbing a toe. Similarly, upsetting news may bring us to tears, causing us to breathe irregularly, whimper, and even fall to our knees.

David Lidov (2005) describes a bodily gesture as “a molar unit of motion, initiated by a single impulse and accomplishing nothing other than an expression or communication” (132). Bowing to an audience after a performance is a clear example of what Lidov describes. This bowing gesture involves bending downward at the waist and reversing the motion to return to an upright position. This action is molar in that its meaning is not an aggregate of the meanings of each of the individual component parts. It is indivisible in that the meaning of this “thank you” gesture is not atomistic; it is not partly found in the downward motion of the upper body and partly found in the upward motion. Instead, we understand the individual actions as coming together through a single impulse to create a unitary gestalt—a single somatic gesture.

Lidov’s definition of the physical gesture appropriately describes the bowing; however, it becomes more useful if we expand it slightly, first by recognizing the importance of the gestural performer, namely, the person enacting the gesture. Although we may take it for granted that a performer is necessary for the realization of a gesture, this element should not be considered unremarkable. *It is a gesture’s implication of embodiment that distinguishes it from other signs.* For example, the green glow of a traffic light is a sign but not a gesture because it is not produced by a sentient being—a gesturer. The traffic light lacks the “impulse” that Lidov requires and, furthermore, does not have an attachment to a volitional entity, which is an essential component of a gesture.

Gestures may also be static bodily positions in which movement is more or less absent. For example, I may stick my fingers in my ears to indicate that I do not want to

hear what someone is saying. Of course, it is necessary to *move* my hands up to my ears to enact the gesture; however, the gesture itself is in the static pose, not the motion toward my ears. In adding to Lidov's definition, we might better describe a physical gesture as *an expressive sign that is produced by a sentient performer through indivisible acts of bodily movement or positioning.*

Having established a clear definition of physical gesture, we are ready to explore gestures that we hear in music. To be sure, "gesture" has been used to describe musical units long before the existence of a developed theory of musical gesture. For instance, the harmonic progression V – vi is sometimes referred to as a deceptive harmonic gesture. Yet without further description of its physical derivation, this progression might better be considered a *metaphor* for a gesture in the same way that picking up the check while dining with close friends is metaphorically a gesture of kindness. Unpacking these metaphors likely reveals a connection to physical gestures; however, the gestures examined in this dissertation are limited to those which have clear physical basis with interpretive consequences.

Robert S. Hatten (2004: 95) succinctly characterizes a physical gesture as a "significant energetic shaping through time," and holds that the sonic analogue—the musical gesture—is achieved through intermodal listening (97). In other words, we imagine that music represents the corporeality of an enacted gesture or the visual appearance of a gesture. Through this process of multisensory integration, these distinct sensory modalities are linked, enabling us to hear a musical gesture that *sounds like* a

particular physical gesture. We might hear the repetition of a short series of fast notes as implying emphatic knocking, pointing, or some other agitated physical gesture because the music—both its mode of production and our conceptualization of the sound—shares the same urgency that is present in enacting and observing the analogous physical gesture.

Scholarly Precedents to a Theory of Gesture

In what immediately follows, I provide a brief review of the literature from the second half of the twentieth century to the beginning of the twenty-first century that positions us to discuss a theory of gesture. Because my theory of gesture relies on agency, narrative, metaphor, and embodiment, I trace the development of these areas as they contribute to gestural approaches to music. My characterization of the development of gesture theory begins with speech-action theory of the 1960s, which leads to the development of agency in music. Once the agent has been theorized and established as a sort of character within a musical work, theorists further explore the way in which these agents interact throughout the unfolding of a piece, paving the path for narrative theories to enter the discourse. At the same, theories of embodiment and metaphor are developed in linguistics and psychology, most notably by George Lakoff and Mark Johnson (Lakoff 1987, Johnson 1987, Lakoff and Johnson 1980, 1999). This orientation suggests that the affordances of the physical human body directly affect our thinking, including the way in which we understand music. By hearing music as bodily movement, scholars such as

David Lidov and Robert S. Hatten have been able to develop theories of musical gesture, that is, hearing music as imitative of physical human movement.

J. L. Austin explores the entailments of human speech as it relates to performative action in his highly influential *How to Do Things with Words* (1962). In this work, he divides speech into two categories: *constative* utterances and *performative* utterances. *Constatives* are those statements which are descriptive in that they serve to relay facts from the speaker to his or her audience. For instance, “Nicholas Sarkozy was the president of France in 2010.” Such a statement is constative because it reports to the reader something factual. We can judge this statement as being “true” because Sarkozy was the president of France in 2010; it is a verifiable fact. If I had claimed that Barack Obama was president of France in 2010, we would most certainly describe the statement as “false” because we know that Obama has never held this office.

Performatives differ in that they “are doing something as opposed to just saying something” (Austin 1962: 133). For instance, when the chairperson says to the board, “This meeting is adjourned,” he or she does not describe or report information. Instead, this utterance *performs* the action of bringing the meeting to a conclusion. Moreover, we do not describe this statement in terms of its truth or falseness as we do when evaluating the above statement about President Sarkozy. We analyze performatives in terms of their appropriateness or felicity based on the context of the speaker and the conditions surrounding the utterance. If a low-level associate were to stand up and declare the meeting adjourned in the same way, we would not be inclined to evaluate the statement as true or false as much as we would deem it infelicitous or unfitting.

The connection between speech and action is important to a theory of gesture because the source of musical expression is often unclear. Where do the expression and emotion in a piece of music come from? To whom do they belong? At one level, the composer is considered the source of the music and its impact on the social and cultural milieu. At another level, the performer is the source. Additionally, we have the conductor, editor, and publisher to consider as well. Edward T. Cone helps to solve this problem in his text, *The Composer's Voice* (1974). Cone explores the relationship between the composer and the unfolding of a musical work. More exactly he suggests that in music the composer takes on a virtual musical persona, a sort of character that he or she enacts while presenting the music to the listener. As Cone argues, we cannot divorce the piece from its compositional source any more than we can remove the speaker from the utterance in everyday speech. The two are inextricably bound together.

Cone's composer, however, differs from the types of speakers that J. L. Austin explores in his speech-action theory. For Cone, art seems to place a sort of mask between the composer and the work. This mask creates a degree of separation that prohibits a work from being "simple autobiography" (1974: 2). Cone sees the relation to literature:

The hypothesis underlying this point of view is an appealing one: that a basic act of dramatic impersonation underlies all poetry, all fiction, indeed all literature worth the name. (2)

This "impersonation" in works of fiction is the basis of the persona that the composer takes on in a piece of music. Because all utterances—linguistic, musical, or otherwise—

imply an utterer, Cone's theoretical persona permits us to hear a unified entity as directing the piece and its unfolding.¹

Cone also discusses virtual characters that can exist within the work itself. He terms these "virtual agents" (1974: 89), a concept that is paramount to a theory of embodied gesture. These virtual agents are embodied through voiceless instruments, acting much like the characters in an opera or oratorio. Cone differentiates the persona—the composer's voice—from the virtual agent in that the former can contain multiple forms of the latter. He writes:

Often a single instrument—whether a solo or a member of a group—is responsible for a number of melodic lines or musical components. In this case, the unitary agent's part, like that of a complex instrumental persona, embraces a number of subsidiary roles. Each of these can be construed as implying its own agent. (98)

With this, Cone implies that there is a nesting of virtual entities.² The persona that we associate with the mask-wearing composer reveals virtual agents, essentially characters within the musical discourse.

¹ Seth Monahan (2013: 329) uses the term "fictional composer" to describe "the person postulated by the analyst as the controlling, intending author of the musical text." For Robert S. Hatten (2004: 231) this agent possesses "super-subjectivity" akin to that of a narrator who presents the musical discourse to the listener in sequence. It is as though this composer agent (or persona) uses the music strategically to tell a story, perhaps one filled with suspense, all the while knowing the outcome. I discuss this further in Chapter 4.

² Monahan describes these as "individuated elements" (2013).

The unitary cohesion of Cone's agents is essential to his theory. Above, we read that Cone differentiates permanent agents that persist throughout the work and temporary agents that appear only briefly before being diffused (1974: 89). He also allows for multiple agents to coexist in the same way that multiple actors exist within a drama (98). These two elements of his theory bring about questions of agential cohesion and agential delineation. Put more plainly, they have us asking, "Which musical elements hold agents together, and which elements delineate agents, keeping them as distinct entities?"

Leonard B. Meyer writes about the "strength of perceived conformance" in his *Explaining Music* (1973: 44-79). Meyer devises a provisional equation to determine which musical elements contribute to the cohesion of a musical idea and which detract from it (Figure 1.1). Having distinct virtual agents within the musical discourse enables a clearer division of gestures and gesturers to be understood. Elements that strengthen the development of a unitary agent are "regularity of pattern," "individuality of profile," and "similarity of patterning" (49). Features that detract from this cohesion and contribute to the dilution of agency are the "variety of intervening events" and the "temporal distance between events" (49). This makes sense because it requires that an agent have a differentiated identity that persists over time. When this identity is not preserved for a long duration, we are less likely to hear the connection across the temporal gap.³

³ Michelle Clater (2009: 67) has developed Meyer's formula to track agential growth throughout a work.

$$\text{Strength of perceived conformance} = \frac{\text{regularity of pattern (schemata)} \cdot \text{individuality of profile} \cdot \text{similarity of patterning}}{\text{variety of intervening events} \cdot \text{temporal distance between events}}$$

Figure 1.1. Meyer’s equation for strength of perceived conformance (1973: 49)

Fred Everett Maus’s oft-cited article “Music as Narrative” (1991) advances agency a step further. He equates the way in which an agent develops over time (i.e., throughout a piece of music) with the development of a character within a drama or a novel. He writes:

Instrumental music consists of a series of events, and the easiest anthropomorphism is to treat these events as behaviors, as actions. Once one begins thinking of musical sounds as actions, rather than just events, the notion of plot or narrative is close at hand. (7)

With this, Maus suggests that agents are not simply discrete units of musical sounds that are lumped together in the mind. We instead hear them as living beings that undergo changes similar to humans. Imbued with a degree of permanence (identity), change, and volition, Maus is poised to develop a theory in which music has a pervasive narrative thread.

Carolyn Abbate’s *Unsung Voices* (1991) echoes many of Maus’s assertions about agency. She challenges Cone’s top-down approach in which the composer injects himself or herself into the music, acting as a puppet master who controls the actions of the agents. Abbate writes:

To Cone's monologic and controlling "composer's voice,"
I prefer an aural vision of music animated by multiple,
decentered voices localized in several invisible bodies. This
vision proposes an interpretation of music shaped by
prosopopoeia [or putting faces on and ascribing human
qualities to nonhuman forms]. (13)

Abbate's bottom-up approach places emphasis on the human-like agents rather than the composer. This has the consequence of privileging the role of the body in music, thus contributing to a gestural approach. From this we can understand virtual agents as not just acting and thinking like humans, but possessing virtual bodies with the same affordances as our own human bodies.

It is certainly no coincidence that this bodily approach to music emerges around the same time that George Lakoff and Mark Johnson begin making a significant impact in cognitive linguistics. Their seminal collaboration, *Metaphors We Live By* (1980), examines the relationship between our language and the way that we perceive the world. In one revealing passage, Lakoff and Johnson write:

[M]ost of the conceptual structure of a natural language is
metaphorical in nature. The conceptual structure is
grounded in physical and cultural experience, as are the
conventional metaphors. Meaning, therefore, is never
disembodied or objective and is always grounded in the
acquisition and use of a conceptual system. (196-97)

With this, Lakoff and Johnson contend that our conceptualization of the world occurs only through the channels of our bodily systems of sensation. We do not understand our environment objectively but rather through the body in which we live. In *Philosophy in the Flesh* they state their approach even more clearly: “The mind is inherently embodied” (1999: 3).

David Lidov’s work proceeds directly from this notion of embodied experience. In his article, “Mind and Body in Music,” Lidov notes the relationship between the words we use to describe music and human bodily movement.⁴ He cites “swing” and *Allegro* as musical descriptors that are borrowed from language used to describe physical movement (2005: 145). Lidov highlights these associations to draw out the physicality that is essential to both the composition and conceptualization of music. More specifically, Lidov interprets most musical movement as a metaphor for physical human movement. Robert S. Hatten furthers this theory by presenting a rich theory of musical gesture in *Interpreting Musical Gestures, Topics, and Tropes* (2004). Hatten describes a musical gesture as a “significant energetic shaping through time” (95). In offering such a succinct and inclusive definition, it is clear that Hatten hears gestures, thus human movement, almost everywhere in music. Hatten deeply explores the entailments of such a perspective, discussing the ways in which gestures contribute to agency and narrative. In his more recent work on agency, Hatten (2012) explores inferences that can be drawn from the interaction of agential gestures and virtual forces in music. He examines how

⁴ “Mind and Body in Music” was originally published in 1987 in *Semiotica* 66, no. 1/3: 69-97. It was later published as Chapter 8 (pp. 145-64) of Lidov’s *Is Language a Music? Writings on Musical Form and Signification* (Bloomington: Indiana University Press, 2005).

the constraints of a virtual environment shape the gestures of an agent, thus influencing its identity and its relationship to the musical discourse. The relationship between Hatten's virtual agent and the discourse can vary widely; an agent can exist as a character within the discourse, a narrator that sequences events, or as an observer that provides commentary about the on-going narrative. With this, Hatten seamlessly sews together theories of gesture, agency, narrative, and emotion, which comprise decades of speculative models of musical expression.

In this section, I have discussed some of the scholarly literature that has led us to a theory of gesture. Below, I review the most current scholarship in gesture theory and discuss some of the gaps that my theory of indexical gestures seeks to fill.

The Present State of Gesture Theory

The connection between musical movement and real movement is so often taken for granted that we sometimes forget that in using the word *movement*, we are metaphorically ascribing both motion and materiality to an otherwise psycho-acoustic phenomenon. Sound, after all, is merely composed of pulsations of air that are received in the ear and transmitted to the brain, where it organizes the aural signals and interprets them within a musical system. In this way, music is an emergent phenomenon. This would seem to suggest that music exists only within the mind and that it is understood metaphorically as motion only for the purposes of verbalized description. However,

Arnie Cox (2001, 2011) proposes a “mimetic hypothesis,” whereby our conception of music as physical movement is not just a creative conceit but rather a consequence of an unconscious imitation. Cox writes:

The initial premise of the hypothesis is that part of how we comprehend music is by way of a kind of physical empathy that involves imagining making the sounds we are listening to. This is a special case of the general human proclivity to understand one another via imitation, which we can refer to as mimetic cognition or mimetic comprehension [. . .]; hence, the “mimetic hypothesis.” (2011: [3])

Cox goes on to discuss *mirror neurons*—a neuron that emits a neural impulse both when an action is performed and when it is merely observed (2011: [13]). Although this neural mimesis (imitation) has been primarily observed in monkeys, a recent study has found the much anticipated and predicted evidence of their existence in humans (Mukamel, et. al. 2010: 750-56). This provides an empirical basis for the mimetic hypothesis; not only are we inclined to conceptualize music that we hear as though we are imitating its production, but we actually cannot avoid it.

Steve Larson has also explored our understanding of music as physical motion. He presents a theory of “musical forces” in his aptly titled book, *Musical Forces: Motion, Metaphor, and Meaning in Music* (2012). This publication follows a series of articles over the course of his career (most notably 1997-97, 2002, 2004) that discuss metaphors that we use to describe musical motion and predict melodic patterns. In *Musical Forces*,

Larson focuses on three forces whose effects are heard in melodic motion: gravity, magnetism, and inertia. We use gravity to describe the reason that a melody descends; magnetism explains the attraction of a melody to a stable pitch; and inertia is the property of a melody that resists a change in direction and speed (2012: 83-100). These, of course, are not forces in a real sense, because a melody lacks physical substance. Instead, Larson's forces are metaphors for the influences that appear to affect the way in which a melody progresses.

Larson's theory of forces provides us with an understanding of musical motion by drawing a connection between the way in which forces act upon objects that we observe and our own bodies. He stands behind a firmly embodied conception of music, writing, "Music *moves*. And something in the way it moves *moves* us" (2012: 61). However, his theory of embodied experience is largely focused on passive movement, in that it primarily concerns itself with the way we *are moved*, not the way we move ourselves. His theory seems to be void of agency in the manner that we typically understand it. If gravity, magnetism, and inertia possess agency, it is a type of non-sentient agency in which these forces act in predictable, prescribed ways. Gravity always causes objects to descend at a predictable rate; the magnetism of the tonic always draws a melodic line toward it, or at least slows its departure from it.

Hatten (2004, 2012) characterizes gravity, magnetism, and inertia as "environmental forces" and notes that they cannot explain melodic leaps or other melodic movements that act contrary to the musical forces. Hatten writes:

If, as I have suggested, we imagine a virtual environment in which Larson's three forces constitute environmental constraints, then such a leap would require additional energy, and the requisite energy cannot be provided by the three musical forces if it contradicts each of them. (2012 [5])

Hatten's argument seems appropriate when we consider the motion of a large musical work. Like a freely bouncing ball, the melodic motion of a piano sonata should stop after a few moments because of the pull of gravity and the diffusion of its energy. Hatten, however, contends that melodic leaps and other motions that result from non-environmental forces imply an agent who continually infuses the music with energy. Thus, we can see Larson's forces and Hatten's virtual agency as complementary. Whichever motions environmental forces cannot explain, the energy of an agent can, and *vice versa*.

Theories of forces and gestures fit together well with Cox's mimetic hypothesis because of intermodality, which Hatten describes as "the capacity for analogous representation across all the senses and motor systems" (2004: 97). According to the mimetic hypothesis, when we encounter a melody, we imagine ourselves enacting a physical gesture with the same gestural contours. An intermodal competency permits us to hear a musical gesture as an embodied physical gesture, and furthermore, we can sense the influence of environmental musical forces as we embody these musical gestures.

Having reviewed the scholarly context within which gesture theory is positioned, we are ready to further explore the entailments briefly described above.

How Agency is Implied

I define an *agent as a unified sentient entity with the capacity for gesture (significant, communicative action)*. Hatten claims that an agent is indispensable for the musical gesture, writing:

When musical events are heard as gestural, then the
implication of agency is inescapable. (2004: 224)

The inference that Hatten draws reflects the notion that gestures are fundamentally embodied signs. Therefore, the mere presence of a communication is enough to imply a communicator, and similarly, the presence of a gesture implies the existence of a *gesturer*.

It would at first seem easy to find the communicator or gesturer in a piece of music, since the performer is enacting a gesture that communicates something to the real listener. For instance, when we observe a violinist bowing, we necessarily see someone engaging in a performance gesture—pushing and pulling the bow across the strings. It is impossible visually to observe the gesture of bowing without simultaneously observing the performer of that gesture because the two are co-material. But as interpreters, this is not the gesture of inquiry. Instead of the *actual* gesture of a performer, we are after the *virtual* physical gesture that is presented by the performer in the (audible) musical gesture. Consider again the first bars of Beethoven's Symphony No. 5. Because the

musical gesture is enacted through the physical performance gestures of string players, we see that the performance gestures imply the presence of these performers. In the very same way, the virtual physical gesture (i.e., knocking) that we hear in the music implies a virtual agent (i.e., Fate). I will further define the *virtual agent as a unified sentient entity with the capacity for virtual gesture*. (I will simply refer to an “agent” when the distinction between a virtual and non-virtual agent is not relevant, or when the more general term is warranted.)

Positing agency within a work is one way to hear the boundaries and unity of a musical idea, such as a motive. Lawrence Zbikowski describes the way in which several characteristics of a motive come together to create a unified, molar identity (2002: 42-48). In his own exploration of Symphony No. 5, he details how some characteristics of the first iteration of the motif change in subsequent entrances while others are preserved. The characteristics that remain constant are those that are most fundamental to the motive’s identity, namely the short-short-short-long rhythmic figure. In preserving the characteristic elements of the figure, we are able to hear subsequent iterations as developments of the initial motive and not the entry of an entirely new idea.

The Implication of the Diegesis

The recognition of gestures “within” music implies the existence of an interior virtual world, and it is in this space that virtual agents enact the movements that listeners outside this virtual world identify as gestures. Philosopher and playwright Denis Diderot describes the distinct separation between the characters of a theatrical drama and the audience sitting in the playhouse. He suggests that these two elements are divided by an invisible barrier or “a vast wall” (*un grand mur*). In an oft-quoted passage from *Discours sur la poésie dramatique* (1771), Diderot writes:

Imaginez sur le bord du théâtre un grand mur qui vous sépare du parterre. Jouez comme si la toile ne se levoit pas.

(Imagine, at the edge of the stage, a vast wall that separates you [the actor] from the audience. Perform as though the curtain did not rise. (259, translation mine)

This virtual barrier between the audience and the actors has become known as *the fourth wall*, in reference to the three walls that comprise a typical box set. The fourth wall acts as an implied boundary; however, unlike the other three, the fourth wall is transparent, allowing the audience to spy on the virtual (fictional) world that the play depicts.

Music is similar to theatre in that it, too, has the potential to represent an interior world that is distinct from the real world occupied by the audience. We consider this virtual world to be the *diegesis* and its characters and events to be *diegetic*. For example, in Dukas’s *The Sorcerer’s Apprentice*, music is used to construct a virtual world containing the sorcerer, his apprentice, and an enchanted broom, all of which are

diegetic.⁵ These characters are virtual agents in that they exhibit sentience, as conveyed by in the actions and gestures that they perform. As noted above, Hatten (2004: 115-16; 2012) describes this diegetic space as a *virtual environment*, observing that it contains virtual agents, virtual forces (Larson 2012), and a virtual terrain. In Dukas's symphonic poem, the virtual environment includes the pails of water that the broom carries and the implied physical space that contains the characters. Although the music does not articulately describe this space, spatial relationships among the three agents are clearly represented. The presence and absence of the sorcerer's musical motif implies his presence and absence within the local space of the work and his distance from his self-aggrandizing apprentice.

By contrast, the elements that lie outside the diegetic virtual world that the work depicts are considered *non-diegetic* or *extradiegetic*. This space is best described in negative terms as anything that is not part of the diegetic space depicted by the work. The most obvious example of the non-diegetic space associated with a theatrical production is the area surrounding the boxed theatre, which most notably includes the audience. In a concert performance of *The Sorcerer's Apprentice*, we would also consider the musicians to be part of the non-diegetic space because they are detached from that which the music represents. They occupy a different conceptual space that is further distinguished by the differences between their performance gestures (e.g., moving the bow) and those implied by the music (e.g., knocking at the door).

⁵ Abbate (1991) analyzes this work with respect to "voice" and potential narrativity.

It is worthy of note that music can also be non-diegetic, with respect to another diegesis comprised of actual actors in a fictional drama; in fact, its role in cinema and staged musicals regularly switches between these two spatial modes. In film, the audience often hears music that the characters within the diegesis cannot hear and are not intended to hear. Although the music plays an essential part of the cinematic production, shading the mood of the scene, it is not always part of diegetic space in which the characters exist. Theatrical characters can also switch their spatial occupancy and often do so in marked contexts, such as in an aside where a character “breaks the fourth wall” and speaks directly to the audience. However, in the repertoire that we examine in this dissertation, it is not necessary to consider such exceptional cases, because the distinction between diegetic and non-diegetic is generally clear.

It is also prudent to clarify that the use of the term *diegesis* in this theory is distinct from its use by classical philosophers. In the *Republic*, Plato—speaking through a Socratic dialogue—places diegesis in opposition with *mimesis* when describing narrative modes (Bk. 3, 392c-394c). For Plato, diegesis refers to a narrator that mediates the space between the exterior world of the audience and the interior world of the drama. This narrator acts as a conduit through which information from within the interior sphere of the works is passed along to the audience. Mimesis, by contrast, occurs when the narrative voice assumes the role of a character. In this mode, the narrator portrays the actions of the virtual world by imitating them, becoming one with a character or number of characters. This conception of mimesis aligns with our understanding of the term, particularly with regard to the mimetic hypothesis put forth by Arnie Cox (2011). But

confusingly, Plato's understanding of diegetic narrative is almost oppositional to our definition above. He describes the diegetic narrative mode as placing a distance between the narrator and the virtual environment; whereas the modern use of the term would typically describe this narrator as non-diegetic in that it observes from the outside, at a remove from the characters.

To summarize, the theory developed in this dissertation does not structure diegesis and mimesis as oppositions. Instead, I reserve "diegesis" for describing the virtual world constructed through musical representation that is distinct from the world of the listener. I use "mimesis" in discussing the way in which music represents physical gestures through listener-perceived similarity with sonic gestures.

A New Theory of Gesture and Agential Perspective

It is interesting that while the embodied musical experience promotes subjectivity, Larson's theory of forces is largely an objectivist one. He describes the predictable influences of his musical forces, implying that they are constants in both the virtual environment he creates and in the way in which we perceive them (2012: 113). But what role does perspective play? What happens when a virtual agent observes the motion of another entity (a virtual agent or virtual object) within the diegesis?

I first considered this while playing a game of catch with my father during a recent summer visit. As the ball left my father's hand, I noticed that it did not appear to move very quickly; however, as it approached me, it seemed to speed up and required me to rapidly make last-second adjustments to catch it. In doing so, I made quick movements

with my arm, hand, head, neck, and eyes so that I could see the ball go into my glove. I did this because the ball's movement *in my field of vision* increased as it approached. Objectively, the force of the ball on its near-linear path remained more or less constant, but from *my* vantage point, it appeared to increase in speed upon arrival. In other words, although the influence of forces on the ball remained the same, their indirect influence on my sensorimotor system was unique to my perspective. Just as perspective causes a change in the apparent influence of real forces on an object, we can say that *agential perspective* causes a change in the way musical forces are perceived. This is better understood for music with the causation reversed—a change in the apparent influence of musical forces can lead us to infer a virtual agent who views an object from a fixed perspective.

Embodying an agential perspective suggests that we are embodying an agent, not an insentient object. In imagining ourselves catching a ball, we take on the perspective of the catcher and mentally simulate the necessary movements to receive the ball efficiently. According to the mimetic hypothesis, listeners are given to interpret musical gestures as *somatic* gestures—movements of a humanoid agent. Thus, hearing and mentally simulating gestures in music implies that we are taking on the perspective of an agent, not of some insensate object.

That we do not embody the object (e.g., the ball) when imagining a game of catch does not present a problem for interpretation because the object is present through inferences drawn from motor movements; that is, it can be inferred from the gestures of an agent. The well-known “hidden ball trick” in baseball is evidence of this. With a

runner on first base, the pitcher throws the ball to the first baseman in an attempt to catch the runner before he can retreat from his lead-off. The trickery comes when the first baseman feigns a throw to the pitcher who pretends to catch the ball. With his eyes trained on the pitcher, the runner believes the ball has been returned to the mound. He is summarily tagged out by the first baseman as he steps away from the base to take a lead-off. In this situation, the runner assumes that the pitcher has the ball because his gestures implied that he has caught it. Music in the same way can imply a *virtual object*, which I define as *an insensate diegetic entity inferred through the gestures of an agent which are unique to its perspective*.

Chopin's Prelude in B Minor, Op. 28, no. 6 offers us an example in which a melody depicts physical motion that implies an observer's perspective.⁶ This short work is one of the composer's more dour settings in the genre. A notable feature of the Prelude is the inverted melody-and-accompaniment texture, shown in Example 1.1; the primary melodic line is situated beneath the accompanimental figuration of the right hand. Hans von Bülow described the right-hand ostinato as imitative of "tolling bells," and its alleged performance at Chopin's own funeral provides evidence of this contemporaneous interpretation (Schonberg 1963: 127).

⁶ Throughout this dissertation, I expound my theory by drawing primarily on music from the nineteenth and early-twentieth centuries because this repertoire has been the focus of much scholarship in gesture and agency. This strong scholarly foundation allows me to develop a new theory of gesture and agency that engages with and complements that of others.

Marion Guck proposes that this work is a depiction of labored breathing (1981: 37-50). The melodic ascent depicts a quick and deep inward breath; it is held for moment; and it is released slowly as the melody descends. This aligns with the solemnity identified by von Bülow but furthers it by suggesting that we are embodying the musical movement. Through this metaphor, Guck places us directly within the diegesis of the piece; we *experience* the gesture rather than simply identifying musical features that imply sadness.

This is not the only interpretation that Guck offers. In her early work on the prelude, Guck (1981) prompted listeners with a number of metaphors and asked that they follow the entailments of these in order to determine what each revealed about the music. It is in a follow-up publication that Guck suggests that the left-hand melody can be that of an observer. She writes:

From the beginning, I described the lines rising and falling;
just above, I described how an arch's line ascends, focuses
and curves. *Movement* infuses both[,] the moment I try to
describe them. It might be like the movement of a ball
thrown in the air, or the movement of an observing eye
[. . .]. (1991: 6, emphasis in original)

The interpretation of the observing eye is different from the breathing metaphor because it directs attention outward from the agent; our focus is not on the gesturer but on the thing that it is observing.

In Example 1.1, we see the left-hand melody of mm. 1-2 make a quick ascent from B₂ to D₄. C#₄ prolongs D₄ with as a neighbor-note elaboration, and the melody descends slowing back to B₂ in m. 3, where the gesture restarts. In m. 1, we hear the energy of the initial ascent through the *crescendo* into beat two; in m. 2, we hear the dissipation of energy and the receding of the object into the distance, as the *decrescendo* complements the descent. Larson's force of gravity is apparent in this passage, but in a unique way. Rather than creating an objective, parabolic trajectory, as might be expected, the quick ascent of the melody and its slow descent imply that the object is traveling into the distance. When tracing an object along this path, our gaze is drawn quickly upward at first and then descends slowly as the object travels away from us. Thus, we do not experience the motion of the melody in the sense that we imagine ourselves moving upward and downward with it; instead, we experience the motion through the embodiment of a virtual agent *observing* a virtual object.



Example 1.1. Chopin, Prelude in B Minor, Op. 28, no. 6, mm. 1-2

As listeners, we train our focus on this virtual observing agent, whose attention is similarly trained on a virtual object receding into the distance. (This object is real for the

diegesis of the virtual observing agent, but virtual for us.) The unique shape of the melody mimics not the motion of the virtual object directly, but the bodily movements of a virtual observing agent as it tracks the virtual object traveling into the distance. Just as the gestures of the baseball pitcher were used to mime his catching the ball, the gestures of the virtual observing agent reveal to us the motion of the virtual object.

Indexical Gestures

The gesture that we hear in the melodic arch of the Prelude in B Minor is distinctive in that it draws attention away from gesturing agent and toward another diegetic entity. In gazing at the virtual object, the agent *indexes* it by pointing its eyes at the object as it moves into the distance. C. S. Peirce identified indexicality as a category of sign whereby meaning is projected through a “real reaction with the object denoted” (1998 [1904]: 307). For example, smoke indicates—indexes—fire on account of a causal connection. Because smoke is often the direct result of fire, the two become semiotically linked with the former making known the presence of the latter. The Dow Jones Industrial Average is another example of an index whereby a single number reflects the value of a group of publicly-traded companies in the United States. Determined by a simple equation, this number is revered as an index of economic health because its value changes as a direct result of stock valuation.

Umberto Eco theorizes the Peircean index in *A Theory of Semiotics* (1976: 115-21). He begins by building on linguistic approaches to indexicality, namely the demonstrative pronouns nouns /this/ and /that/. He identifies these words as indexical

based on their contextual dependency for meaning. In other words, /this/ and /that/ earn their meaning according to the context in which they appear. For Eco, the meaning of /this/ merely denotes a closeness to the speaker, while /that/ denotes a comparative distance from the speaker. Within a given context, they are able to transmit significant meaning. Consider the differences in the follows sentences spoken across a table:

- (a) *This* is my apple.
- (b) *That* is your apple.

The apple in (a) is close in proximity to the speaker while the apple in (b) is relatively distant. However, this relationship is reversed if the addressee were to make the very same utterance. In this second instance, the apple initially referred to by the demonstrative /this/ has become known by /that/ in the words of the second speaker.

Demonstrative pronouns are indexical in that they have basic meanings ('near' and 'distant'), but in practice a more specific meaning emerges through context. They are very much akin to pointing (*deixis*); in fact, one can imagine using an index finger while speaking in order to further clarify or emphasize their meaning. Thus, pointing gestures are referred to as deictic signs (as are demonstrative pronouns). Pointing is one of the most basic indexical gestures that we encounter daily. By extending an arm outward from the body and protracting the forefinger, we draw attention to an object that is beyond our reach. Although we perform and respond to this gesture frequently and without thought, Eco's (119) analysis of this mechanism is rather interesting. He finds that pointing typically involves four features: longitude, apicality, movement-toward, and dynamic

stress. When these features come together, a gesture emerges that effectively communicates a connection with a distant object, person, or location.

Longitude and *apicality* are the two most crucial features of the pointing gesture.

With regard to these, Eco writes:

[The pointing gesture] is longer than it is bulky and ‘stops’ at the top of the fingernail. This observation might seem rather obvious, but suffice it to think of an iconic surrogate of the pointing finger (for instance, of the image of an arrow in a road signal) to discover that these features are indispensable. (119)

He continues by observing that the gesture implies an extension beyond the tip of the finger toward the indicated object or location. It should be noted that apicality does not simply refer to the end or terminus of the longitudinal extension; we should remember that the pointing gesture has two ends: the fingertip and the body proper. Eco notes that the apical end is marked (significant) in that it possesses a specific meaning about direction that the other end does not (119). Like the person gesturing with a pointed finger, an arrow in a traffic signal has two ends. On one end we find a narrow tip or arrowhead that the opposite side may lack. The special property of the apical end expresses a unidirectional motion toward the remote object or location that the other end does not convey. In other words, while the longitude of the gesture produces the path, it is the apicality of the fingertip that tells us of the direction.

Eco describes *movement-toward* as a feature of the pointing gesture that can sometimes substitute for longitude. He writes:

[Movement-toward] must always exist, even if imperceptible. In some other kinesic expressions it is absolutely indispensable, as occurs when somebody ‘turns’ his head or glances toward something. In this case the features //movement toward// and //apicality// suffice, and //longitude// is not needed. (119)

I have two criticisms of the above description. First, Eco seems to be refuting his previous claim in which he characterizes longitude—along with apicality—as “indispensable” (119). This oversight, however, should not upset his entire theory. Instead, we might benefit from characterizing longitude and movement-toward as two manifestations of the same principle. Longitude is a characteristic of the static form of the feature while movement-toward is its dynamic form. My second issue with the preceding description of movement-toward is that Eco seems not be describing literal movement that is toward an object. When we turn the head to look at an object, we do not get physically closer to it; we merely change our orientation. It seems that Eco confuses movement-toward with the physical repositioning of the head. Although this gesture undoubtedly points toward the object, it is the apicality of the face and its implied motion of the gaze that achieve this indexicality. We look toward an object, implying a unidirectional connection between the eyes and the distant object.

Dynamic stress is the fourth feature of pointing in Eco's analysis (119-20). By this, he refers to the repetition, urgency, and the force of the gesture. He describes:

When a finger points with comparatively little energy it means <<close>>, and when pointing with greater energy (the gesture being more ample and abundant, with the arm itself participating in 'propelling' the finger) it means <<distance>>. (119)

I often experience this difference on the first day of the semester when a student standing outside of a classroom asks whether he or she has the correct room with a casual point toward the door. If the unsure student is standing outside the wrong classroom, I may explain with an exaggerated pointing gesture that he or she needs to go *way* down the hall. At times when a student is especially lost, I direct them to another building across campus, making a grand gesture in which my arm moves all the way across my body.

It is important to note, however, that dynamic stress is often used to index objects that are very close to us. When sharing a particularly brilliant passage from a text with a colleague, I am known to point repeatedly and forcefully at the page to show my excitement. In these instances, the indexed object—the page—is in the foreground. The dynamic stress that I use to enact the pointing gesture shows my intense feeling about this immediate object. By contrast, I may direct my colleague's attention by flippantly pointing at a page that contains an uninteresting or specious argument. My glib gesture in this situation expresses a metaphorical distance from the thing to which I point. Thus, it seems that the dynamic stress that accompanies a pointing gesture expresses a difference

in the immediacy of an object, but not simply in the way that Eco describes. On one hand, the increase in the gesture's force and repetition convey a distance, while on the other hand, this emphasis expresses importance and cognitive immediacy.

Pointing with an extended arm and index finger is, of course, not the only way to draw attention to something. In his analysis of musical gestures, Hatten (2004: 151-52), draws our attention to a gesture in the final movement of Beethoven's Symphony No. 2 shown in Example 1.2. Hatten writes that as the piece begins "a scurrilous rascal thumbs his nose at us (or makes another appropriate Italian gesture)" (151). Although Hatten minces words, perhaps the "Italian" gesture that he refers to is a hand movement in which the back of the fingers are placed beneath the chin before moving outward with the arm in a quick sweeping motion. The meaning of this widely-known gesture ranges from a simple dismissal to a harsh profanity on a par with the American middle finger. But regardless of the degree of insult that it connotes, the expression shows derision *toward* a person or object.



Example 1.2. Beethoven, Symphony No. 2, Op. 36, iv, mm. 1-2

In analyzing this dismissive gesture, we see that it possesses the same characteristics as the finger-pointing gesture. The offending gesture exemplifies longitude, movement-toward, and dynamic stress. Furthermore, apicality exists at the end of the gesture. The movement-toward is perhaps the most identifiable feature. The motion of the hand from beneath the chin to a position out in front of the body directs attention to the person or object receiving the insult. It is longitudinal in that the length of the arm assists the hand in moving outward. The gesture culminates with the arm almost fully extended. Similar to the pointing gesture, it is the juxtaposed fingertips that give it apicality. Along with movement-toward, the apex of the fingertips at the completion of the gesture are marked, indicating the location of the dismissed person. Dynamic emphasis is found in the quick sweeping motion outward from the body. A rapid and forceful motion toward the target is far more biting than a light and casual extension of the arm.

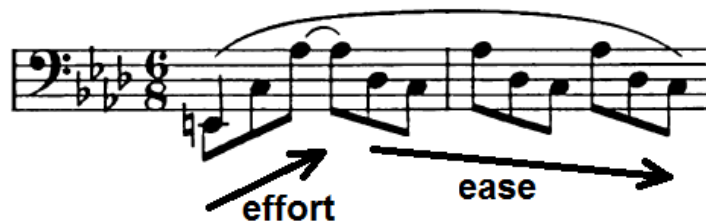
A careful examination reveals the four elements of the pointing gesture in Symphony No. 2. We hear the preparation of the gesture in the two-note anacrusis figure. The slurred F-sharp to G depicts the back of the hand approaching the chin, “cocking” the dismissal before it is fully enacted. After a brief pause, we hear the commencement of the gesture proper. The *sforzando* and trill on C# portray the dynamic stress that adds intensity to the hand motion while longitude is conveyed through the texture. Beethoven scores the opening measures in unison, providing a more directed motion that, like the human arm, is longer than it is wide. As noted above, this projects a directional character to the gesture, one that extends outward from the body and toward the object of insult.

The temporal aspect of the melody assists in creating movement-toward. With dynamic stress enacted by the trill, the melody seems to push toward beat one of m. 2. In this melodic motion, we can hear the hand physically extend outward from the chin to nearly a full extension of the arm. This is the point of release for the gesture, as beat one is a metric strong point. The descent to A that immediately follows depicts the fall of the arm as it fully reveals the palm of the hand and underside of the fingers. This final pitch of the musical gesture also represents the apical point of the corresponding physical gesture which, as a stopping point, helps to give direction to the insult. While the temporal aspect inherent in music would seem enough to provide the unidirectional movement-toward that the apex offers to a static gesture, such as a traffic arrow, one element is not more fundamental than the other. Apical points that segment the music are just as inherent as its temporality. All musical gestures have directed temporal motion in the same way that all gestures have an apical point in which they end. In this example, we see the way that multiple features come together to form the emergent gesture.

The gaze is another type of indexical gesture that is a special form of pointing. The language that we use to describe the act of looking is similar to that of pointing. We *cast* glances *at* people, look *toward* objects, and peer *through* windows. The marked apicality of the face, the craning of the neck, and the straining of facial muscles come together to create a type of pointing gesture that is fundamentally similar to that of the outstretched arm and pointed finger. The causal relationship with the distant object that we index with our gaze is present in the way it affects our visual pointing. A small or distant object might cause the observer to squint; a frightful sight might induce a recoil.

Although the latter action seems to counteract the motion toward that contributes to pointing, our language also suggests that looking involves inward motion as well. We *take in* beautiful scenery. We *get* a glimpse of a celebrity. The mechanism of sight involves looking toward an object and receiving its image as reflected light, and our language for describing vision demonstrates our awareness of this reciprocity.

This two-way path between an observer and an object permits us to infer an embodied vision while revealing information about a distant object. Liszt's *La lugubre gondola* is widely heard as a depiction of Wagner's own funeral procession down the Grand Canal in Venice. The dark character and the accompaniment texture that reflects the strokes of a gondolier's oar both contribute to this reading. In Example 1.3, we see the quick and forceful ascent from E₂ to Ab₃ that embodies the effort that the gondolier must exert to propel the craft through the water. This is followed by three beats of ease as the gondolier circles the oar and prepares for the next stroke.



Example 1.3. Liszt, *La lugubre gondola*, S. 200/1, energetic pattern of the gondolier

We can imagine ourselves standing alongside the river watching the gondola float elegantly down the canal while being overcome with rich, somber emotions. As the boat

approaches us, its visual trace in our field of vision increases in size. We hear this in the *crescendo* that begins in m. 95 (Example 1.4). As the gondola draws near to us, the emotion welling up within our body similarly grows more intense (*rinforzando molto*). It is most immediate in mm. 101-106, where the gondola is largest in our view and the gut-wrenching sadness is most palpable. This is the height of rhythmic activity in the coda. Finally, the watercraft begins to move away from us m. 107 as the dynamics lessen and the sight of it grows smaller. The energy of the rhythm starts to subside, stopping completely in the final four measures. The dynamics become faint as the piece culminates with an almost impossibly soft *ppp* in the final two measures. As the work ends, the gondola has disappeared from view and we are left only with the lingering impression that it has made.

Example 1.4. Liszt, *La lugubre gondola*, S. 200/1, mm. 94-120

As we continue to point our gaze toward the craft, it recedes into the distance. Its size grows ever smaller in our field of vision, just as the volume of the piece diminishes. We have come to know the distance and character of the boat by embodying the sensual

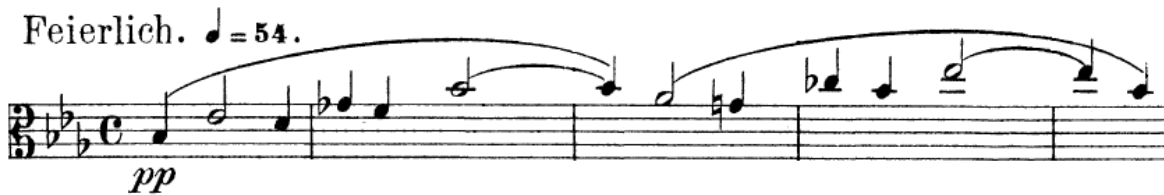
(aural) trace of the music in the same way that an observer along the Grand Canal would embody the sensual (visual) trace of the actual gondola.

The virtual object that the observing agent relates to the listener can also be stationary, with the indexical gestures portraying the physical movements that the observing agent enacts while visually taking in the entirety of the virtual object. For example, the fourth movement of Schumann's Symphony No. 3 in E-flat Major (*Rhenish*) reflects the composer's sight of the Cologne Cathedral during his visit to the Rhineland. Schumann captures the essence of his experience by succinctly marking the score "*Feierlich*," literally "solemn." In listening to this movement, we can imagine ourselves in Schumann's position, observing the grand spectacle of the exterior of the massive structure. Philosopher Charles O. Nussbaum (2007) writes about this movement, noting:

Certain bodily movements, eye movements, head movements, and slow walks round are required to take in such a structure. Models of observer and object located in musical virtual space complement the musical plan. (231)

We can hear these ascending bodily movements of an observer in the principal theme of the movement. Example 1.5 depicts the melodic segment that I interpret as the gradual ascent of the head, neck, and eyes of an observer—perhaps representing Schumann himself—taking in the sight of the cathedral. The overall rise of the melody is noticeably patient, not at all deliberate. The ascending leaps of a fourth are followed by half-step

descents which are mimetic of the awe-filled gaze.⁷ The observer’s head rises as its line of sight scales the front facade to the cathedral’s tympanum where it momentarily pauses to examine the detail of the biblical figures carved in relief. The view then continues its ascent as it traces the pointed lancets before again pausing contemplatively to reflect on further details of the architecture. We can imagine this process continuing all the way up to the spires of the cathedral’s towers.



Example 1.5. Schumann, Symphony No. 3 in E-flat Major, iv, mm. 1-5, first trombone

Nussbaum’s approach to the musical representation seems to be an embodied one in that he posits a virtual observer much like my virtual observing agent to enact the gazing gestures which are then embodied by the listener. Although he uses the virtual observer as a channel through which the cathedral (as a virtual object) is relayed to the listener, he seems to turn his back on this model later in the same paragraph:

[T]he musical listener takes on not only the body sets of the

observer under these conditions, *but also simulates the*

⁷ It is important to notice the way in which the interpretation of the gesture is informed by non-gestural aspects of the work. The slow tempo, the reverent brass, the minor mode, the prominence of fourths, and the imitative polyphony recall sacred genres. By placing the gesture into this context, we are able to better articulate its meaning. This allows us to hear the ascending and descending physical movements as those performed by a virtual agent expressing deference and wonder.

“*bodily sets*” of the observed virtual object, transformed
under the influence of a kind of poetic animism.

(Nussbaum 2007: 231, emphasis in original)

With this, Nussbaum claims that in addition to representing the observer’s view of cathedral, the music can represent the cathedral itself through anthropomorphism.

Although this seems to subvert the observing agent, Nussbaum is actually suggesting that we can hear the cathedral as an agent itself, possessing human features including bodily structure and comportment. I do not disagree with this. Recalling Larson’s (2012) theory of forces, we hear melodic movement as physical motion, and in Hatten’s (2012: [5]) example of melodic leaps, these can be movements of an anthropomorphic agent, even though we do not ascribe specific features to it. Just because Schumann’s program for the fourth movement of *Rhenish* indicates that the virtual object is a cathedral, it does not necessarily mean that we cannot embody it as a virtual human agent and identify directly with it.

However, a problem arises with Nussbaum’s desire to subvert the observing agent and claim that the cathedral can represent itself directly to the listener through mimesis as a result of his first interpretation of the movement. The observing agent observes the cathedral from a fixed perspective. The tower spire is considered *high* because it is *high relative to the observing agent*. The principal melody ascends because it reflects the physical efforts of the observing agent in capturing the trace of the cathedral in the sensorimotor system. The cathedral is tall but it is not growing taller; a direct representation to the listener would not involve an ascending melody unless it were

becoming taller or changing its anthropomorphic bodily comportment. It seems, however, that we cannot avoid representing the cathedral through height gestures, because its height relative to us is such an important component in our conceptualization of the structure. We think of the cathedral as something that takes effort to observe; thus, the observer is essential to its conceptualization. Following a proposal by Nussbaum (139), if we were to represent a cathedral in a game of charades, we would not get on all fours or contort our bodies into the cross-like shape of the cathedral's floor layout; instead, we would stand tall and draw attention to our height, perhaps raising our arms to make a steeple above our heads. It would seem difficult then to remove the observing agent from the representation of the cathedral in this movement because our observation of it is integral to our understanding of it; its key feature—its height—is a relative quality. In removing the observing agent, Nussbaum is actually just overlooking it; he is forgetting that observation is an essential part of the cathedral as a concept. In examining other works below, we will find that observation is an often neglected aspect of conceptual development, one that is largely overlooked by current theories of agential gesture.

Breathing or Gazing: Lobbying for the Indexical Gesture

The suggestion of indexical gesture begs an important question: Why should we hear gestures as indexical when we can just as well hear them as non-indexical gestures. Posing the question in a specific context, we might ask: In Chopin's Prelude in B Minor,

why should we hear the gesture as gazing when we can hear it more simply as deep breathing?

I am not suggesting that Guck's metaphor of the breathing agent is inappropriate; in fact, I personally think that it provides a visceral contribution to an interpretation that is otherwise vague. The symbolism of the tolling bells and connotations of the minor mode set the mood, but the breathing metaphor makes it more descriptive through personal engagement. The music suggests the heaving breathing of a virtual agent, and the listener co-experiences this through active or passive (imaginative) mimesis. In other words, the listener actually enacts the breathing of the virtual agent or simply imagines this breathing.

Hearing the prelude's opening gesture as indexical, however, adds another element to the interpretation of the piece *by giving us access to other virtual entities within the musical diegesis*. The gaze indexes a virtual object traveling into the distance that we can understand as something that draws our focus away from the present moment toward a temporally distant place. As a result, the gaze may allegorically represent a pensive look into the past or into the future, or perhaps both simultaneously.

This is not irreconcilable with Guck's breathing metaphor because the musical gesture can be a *multivalent signifier—a sign with multiple meanings*. The gesture can represent the gaze of an agent *and* at the same time represent this agent's deep breathing. In this way, the gesture is similar to a *pun*, in which the two meanings of a single word are exploited to enable a sentence to allow for two readings. A linguistic example from

Shakespeare's *Romeo and Juliet* (3.1.65-66) is illustrative. Mercutio, just having been stabbed, utters:

Ask for me tomorrow, and you shall find me a grave man.

The play on words is obvious:

meaning a: Ask for me tomorrow, and you shall find me a *serious* man.

meaning b: Ask for me tomorrow, and you shall find me *in a tomb*.

We see here that /grave/ is a multivalent word. Interpreted as in (a), Mercutio expresses the seriousness of the wound. Interpreted as in (b), Mercutio specifically acknowledges that the wound is fatal. Although /grave/ means *serious* in (a) but means *tomb* or *cemetery plot* in (b), the two meanings complement one another to enrich the expression.

The dual meaning of the gesture in the prelude functions similarly to the *pun* presented above. The quick-ascent-slow-descent melodic arch that unfolds across the first two measures can be heard as both a representation of breathing and gazing simultaneously. In other words, this one musical gesture says two things about the agential subject that we might actually synthesize into a more complex physical gesture. One can imagine breathing deeply *while* gazing pensively toward the far-off horizon. In joining these interpretations, we create an even more vivid image of the agent's actions and emotional state.

Conclusion

The relationship between the mind and the body exists at the foundation of human cognition. As a result, our conceptualization of the world around us is necessarily

mediated by the body and more specifically its channels of sensation. Because we are drawn to understand the world in terms of our own bodies, it is not difficult to understand why we ascribe human characteristics to non-human entities. Chairs have legs. Clocks have hands. Similarly, we describe changes in state as movement that we attribute to an agent. Gravity pulls objects down. Lightning is the scourge of Zeus. Humanity is the result of intelligent design.

Our understanding of music is no different. Considering language that describes head-motifs and musical movements, the leap to agency and gesture appears to be a natural step. Taking yet another step, we approach a theory of indexical gesture that reveals a virtual observing agent who can serve as the lens through which we view the diegetic, virtual world that the music constructs. This theory of the virtual observing agent does not require a wholly new reconfiguration of our musical understanding. Instead, it develops ways in which we already conceptualize music, while enhancing our understanding of music's expressiveness.

CHAPTER 2
MATHEMATICAL AND LOGICAL FORMULATION
OF INDEXICAL GESTURES

Introduction

We do not earn the virtual observing agent easily. A theory of musical gesture begins with human embodiment (Lidov 2005, Cox 2010). Theorists such as Cumming (2000) and Hatten (2004, 2012) have moved from an actual listening body, moving to the music, as in Cox's early mimetic hypothesis (1999), to a virtual entity that implies a virtual human agency. Studies of intermodality, implied corporeality, and virtual agency have led to the construction of a strong foundation for interpreting human gestures in music. Based on this work, I have developed a theory of particular *indexical* musical gestures that imply a virtual observing agent. Here, I present mathematical and logical constructions to support these indexical gestures, and hence an implied virtual observing agency. I follow up on my previous analysis of Chopin's Prelude in B Minor, Op. 28, no. 6, demonstrating that the thematic melodic arch implies a gazing gesture and therefore a virtual observing agent. In the second part of this chapter, I track the chain of inferences that a listener makes when hearing the presence of a virtual observing agent in music. This is illuminating because it not only helps to support a theory of virtual observing agency, but it also articulates the inferential path that one follows when hearing agency, action, and gesture in music.

Part I

Evidence for the Gazing Gesture

Consider the arch-shaped trajectory of an object and its effect on an observer. As discussed previously, I hold that tracking an object, such as a ball, as it travels along a parabolic path, entails certain head, neck, and eye movements that mimic a melodic arch contour with a single climax near its beginning. I explain this tracking gesture using algebra and trigonometry in what follows.¹

I noted in the previous chapter that if we were to hit an object, such as a baseball, up and outward from us, the object would travel along a parabolic path because its horizontal speed remains constant due to inertia, while its vertical speed would be variable on account of gravity. This is a primary tenet of ballistics—the study of the motion of objects that have only an initial moment of propulsion before being subjected to environmental forces. Every parabola can be expressed algebraically as a quadratic equation or second order polynomial.² The simplest quadratic equation is

$$y = x^2$$

Figure 2.1 depicts this equation.

¹ Through personal correspondence, Isaac Gerg, Shannon Kolensky, and Chris Linick contributed to the development of this section by proofreading the formulae. Isaac Gerg provided valuable assistance in helping me to conceptualize the arch and to express the virtual observing agent's movements in terms of a trigonometric function.

² The equation is termed “second order” because the highest power of x is 2.

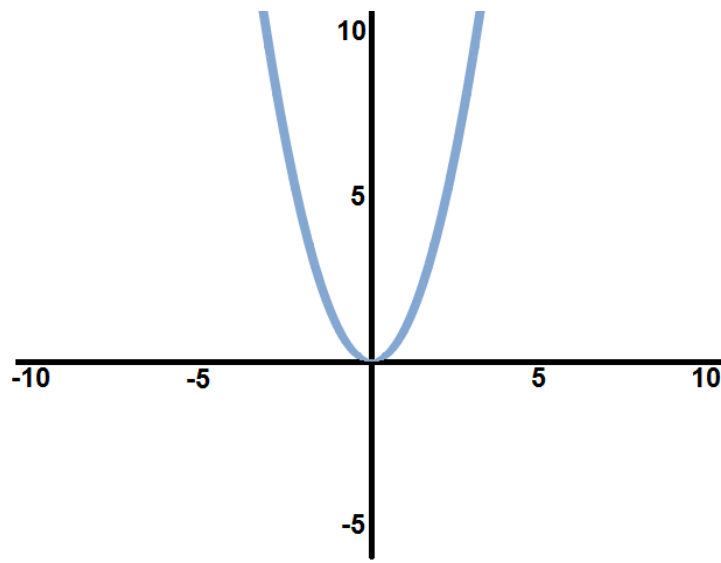


Figure 2.1. Parabola $y = x^2$

To invert the shape so that it reflects the correct parabolic trajectory of the object, we multiply x^2 by -1. This renders the equation

$$y = -x^2$$

In performing this operation, the plot pivots on its apex, which now represents the high point of the object's path of motion. Figure 2.2 illustrates this.

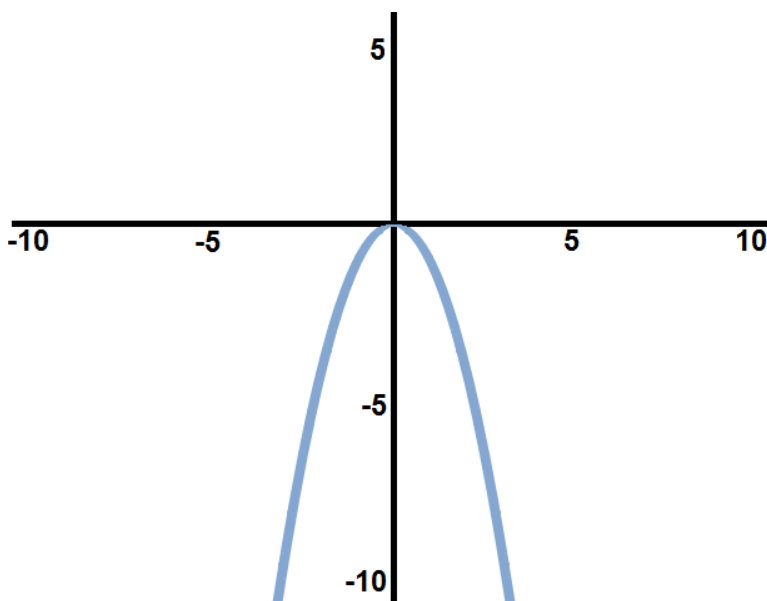


Figure 2.2. Inverted parabola $y = -x^2$

We now have a parabola that looks closer to the trajectory of the object. By manipulating the equation further we can place the inverted parabola in a position on the graph that more clearly reflects a more typical (slightly less steep) path of motion (e.g., roughly a pop fly to second base), with a clearer relationship to home plate (now represented by 0,0) and to the ground (now represented by the x-axis). I have altered the previous equation to take on the following form:

$$y = -0.04x^2 + 4x$$

Multiplying x^2 by a value less than 1 widens the parabola; multiplying x by a value greater than 0 places it in the first (upper-right) quadrant. This equation is depicted in Figure 2.3 with an arrow indicating the motion of the object. In this figure we see the apex (or vertex) of the parabola at point (50, 100) and intersecting the x-axis at (0, 0) and

(0, 100). These intersecting points are known as *x-intercepts*. It is important to note that the equation $y = -0.04x^2 + 4x$ is the only one that yields a parabola with this vertex at (50, 100) and x-intercepts at (0, 0) and (100, 0). If we imagine the distance in terms of feet, then this would be the path of motion for an object that reaches a peak height of 100 feet and lands 100 feet from its initial point of propulsion. It is the force of gravity that slows its motion on its ascent and increases it as it falls. Interestingly, however, gravity does not affect the horizontal velocity of the object. Because it does not slow the movement of the object's horizontal motion, its movement along the x-axis remains constant unless it makes contact with another object. In other words, it continues to move away from the y-axis without slowing down or speeding up. Of course, in real life, wind resistance (aerodynamic drag) would affect the horizontal motion of the object; however, an observer from home plate could not readily detect its negligible effects at the distances proposed in this scenario. Thus, I have removed the effects of drag from this analysis. Figure 2.3 provides a clear image of the object's movement from its point of propulsion (0, 0) to the moment that it returns to the earth.

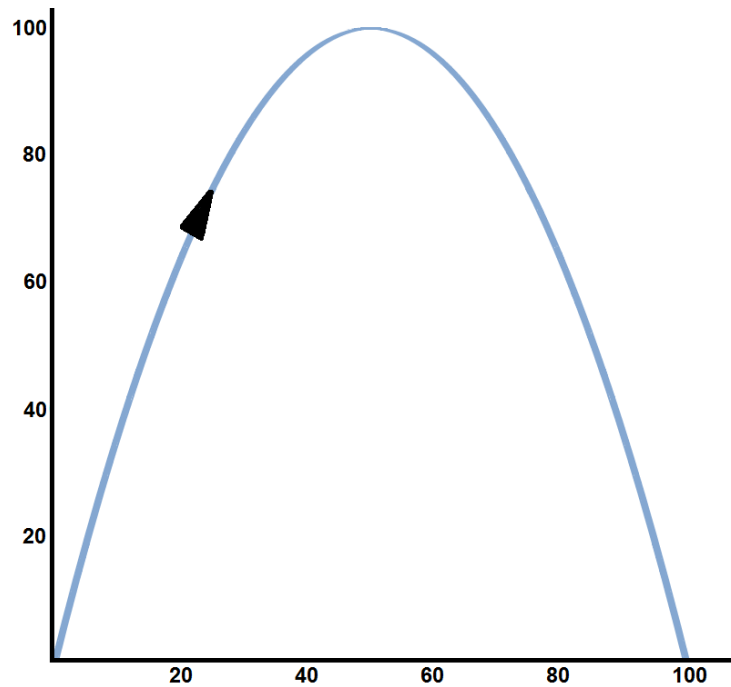


Figure 2.3. Path of motion of object (represented in feet) $y = -0.04x^2 + 4x$

Although an object travels along a parabolic path like the one depicted above, I further intend to map the collective head, neck, and eye movements that an observer must make in order to track the object moving away from the observer positioned at home plate. It is important to remember that the graphs thus far represent an orthogonal perspective in which the object's path of motion is perpendicular to the observer, for example, standing on first base. Yet, the observer whose motions we want to determine is positioned at home plate, or $(0, 0)$ on the graph while tracking the propelled object. From this perspective, the object appears to move *outward* from the observer. I will demonstrate that the contour of the observer's bodily tracking gesture is *not* parabolic,

nor does it possess the symmetry of a quadratic equation. As mentioned in Chapter 1, the head, neck, and eyes would move upward quickly at the beginning of the object's ascent, hang for a moment, and then descend more slowly as the object fell into the distance. Placing my focus on visual tracking, I am now concerned with the *angle of vision* relative to the x-axis (i.e., the ground). The angle of vision is at 0° relative to the x-axis when we are looking at someone of equal height standing in front of us. When at 90°, we are looking straight up at the point on the ceiling just above our heads.

We can calculate the zero-point observer's angle of vision by using trigonometry. If we overlay the parabola with a right triangle, as shown in Figure 2.4, we can then calculate the angle (θ) at each given point on the graph. For example, at its apex (50, 100), the right triangle has legs of 50 feet and 100 feet in length. Trigonometric functions (i.e., *sine*, *cosine*, *tangent*) help us to determine the unknown angles of right triangles. The *tangent* (tan) of a right triangle involves the relationship between x and y to the extent that

$$\tan \theta = \frac{y}{x}$$

Expressing this as an inverse function allows us to solve for the angle in question (θ):

$$\tan^{-1} \frac{y}{x} = \theta$$

Thus, with regard to

$$y = -0.04x^2 + 4x$$

when

$$x = 50$$

$$y = 100$$

$$\tan \theta = \frac{100}{50}$$

$$= 2$$

$$\tan^{-1} 2 = \theta$$

$$\approx 63.4^\circ$$

With this equation, we learn that when the object is at its highest point, the viewing angle of an observer, whose eye is exactly at the point of propulsion (0, 0), is approximately 63.4° above level.

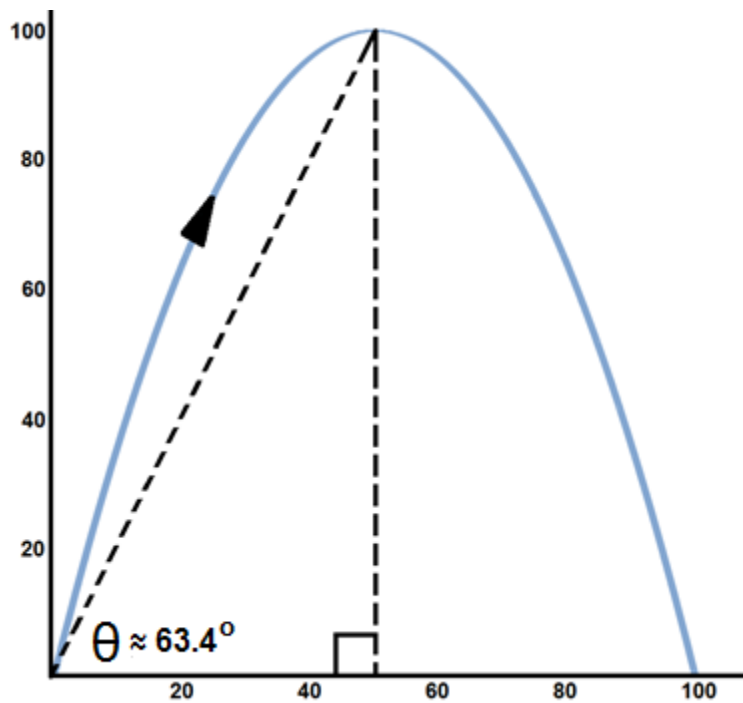


Figure 2.4. Right triangle placed against path of motion of object
 $y = -0.04x^2 + 4x$

This calculation seems to align with our intuition; however, it is impractical to evaluate the viewing angle directly at the point of propulsion (0, 0) because there is always some space between the eye and the object. Although this space may seem negligible as the object travels into the distance, ignoring it would not permit us to accurately portray the viewing angle in the first few moments of the object's motion. Leaving directly from the eye, the object would be almost directly above the head of the observer in the first few inches of travel, forcing the head, neck, and eyes to their most extreme upward positions. However, placing the observer at, say, 10 feet from the point of takeoff provides a more realistic model of observation. We might switch our analogy now from the baseball player hitting a ball to a spectator on the same level, but 10 feet

behind home plate. The following equation preserves the shape and dimensions of the trajectory, merely shifting them 10 feet to the right:

$$y = -0.04x^2 + 4.8x - 44$$

Figure 2.5 plots the above equation.

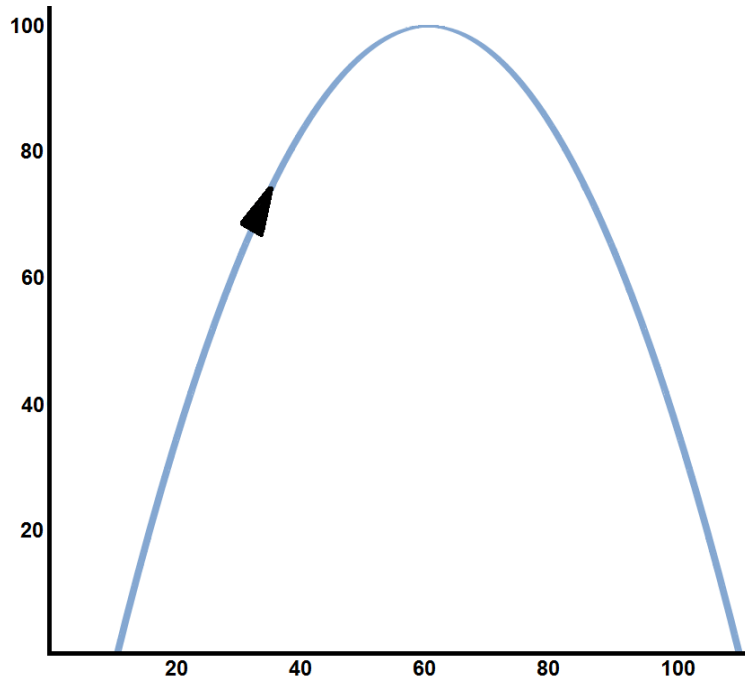


Figure 2.5. Path of motion shifted 10 feet right on the x-axis
 $-0.04x^2 + 4.8x - 44$

With this more feasible point of view, we are ready to plot the viewing angle for all points along the object's path of motion. We take the equation that mapped the object's trajectory—now offset by 10 feet—and apply the tangent function accordingly:

$$y = -0.04x^2 + 4.8x - 44$$

$$\tan \theta = \frac{y}{x}$$

$$\tan^{-1} \frac{y}{x} = \theta$$

$$\tan^{-1} \frac{-0.04x^2 + 4.8x - 44}{x} = \theta$$

Since x represents both distance and time, every value of x gives us the viewing angle (θ) at each moment. This is represented in Figure 2.6. Here, x represents the ground distance of the object in feet. Because this rate of motion is constant, x can also be seen as representing time, but in non-descript units (e.g., $20 \neq 20$ seconds). With respect to the y -axis, each point represents the viewing angle (θ) in degrees; thus θ in the above equation equals y in Figure 2.6. As the object begins its ascent at a point 10 feet away from the observer, the viewing angle rises quickly. It begins to steady as it approaches its widest angle ($\approx 65^\circ$).³ The head, neck, and eyes descend slowly as the object approaches a distance of 110 feet from the observer.

³ Note that the apex of the viewing angle is not the same as the apex of the parabolic movement of the object. The object reaches its highest point at $x = 60$, but the head, neck, and eyes begin descending at about 33 feet.

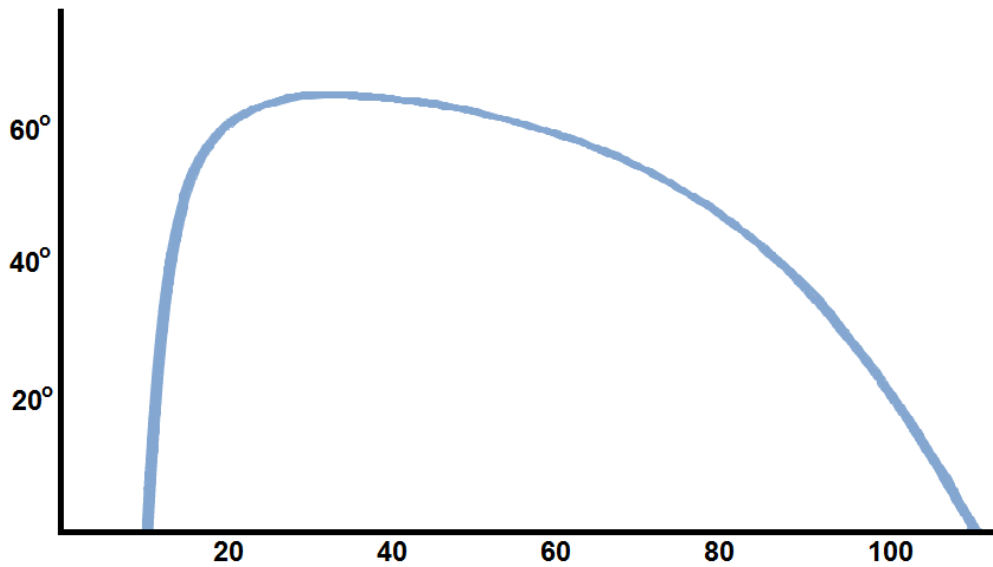


Figure 2.6. Angle of vision at 10 feet from point of propulsion

$$y = \tan^{-1} \frac{-0.04x^2 + 4.8x - 44}{x}$$

Furthermore, we can graph the rate at which the viewing angle changes by finding the derivative of the equation depicted in Figure 2.6. The derivative evaluates change by measuring the slope of a series of *tangent lines*. These are straight lines that have only one point of contact with the plotted equation. If we place a strip of tape around the circumference of a basketball and place a yardstick parallel to that tape, the yardstick represents a tangent line because it only touches the basketball at one point. Placing a straight line on the outside of the plot in Figure 2.6, similarly, creates a tangent line. The derivative calculates change in the slope of the tangent line as it moves across the plot. The tangent line in the ascending portion of the graph has a steep, positive-value slope. At its apex ($x \approx 33$ feet) the tangent line has a slope of zero. As the object falls, the

tangent line slants the other way, producing a negative-value slope. Since the rate that the observer's viewing angle changes is more important than whether the head, neck, and eyes are moving upward or downward in this equation, we calculate the absolute value of the derivative, ensuring that the change in motion is always represented by a positive number:

$$y = \left| D \left(\tan^{-1} \frac{-0.04x^2 + 4.8x - 44}{x} \right) \right|$$

Each point in Figure 2.7 represents the rate at which the head, neck, and eyes move collectively to track the object along its path at any given moment. These plots together represent the acceleration (and deceleration) of these tracking movements over course of the object's path of motion. Notice the observer's movements are quicker at the very beginning of the object's ascent. They become slower as the object nears its highest point. These tracking movements then speed up as the object descends; although, they do not change as quickly as they do during the ascent.⁴

⁴ Figure 2.7 does not specify the value or units of the y-coordinates because a standard unit to measure change of angle does not exist. Each y-value is best understood in relative terms.

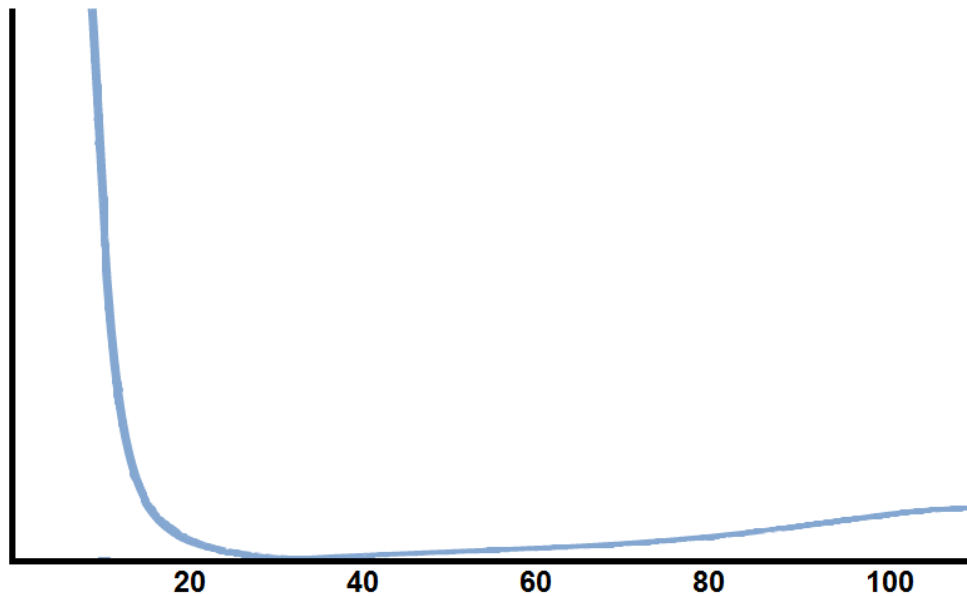


Figure 2.7. Rate of movement in observer's viewing gesture tracking an object moving 100 feet. The y-axis represents the speed of the observer's movements, beginning fast, slowing at the peak, and gradually accelerating as the object descends.

$$y = \left| D \left(\tan^{-1} \frac{-0.04x^2 + 4.8x - 44}{x} \right) \right|$$

In Chapter 1, I claimed that the quickly-ascending-slowly-descending melodic contour of the opening measures of Chopin's Prelude in B Minor, Op. 28, no. 6 represents a gesture of gazing—pointing with the eyes, as it were. The above modeling demonstrates that when an observer tracks an object on a typical parabolic path, she collectively moves her head, neck, and eyes quickly upward before descending at a more measured pace. Figure 2.6 plots the angle of the observation's motion relative to the ground and shows its change as the object of interest travels. The strong similarity between the Figure 2.6 and the left-hand gesture that begins the prelude supports the

musical interpretation of a virtual observing agent gazing at something which has been propelled into the distance.

Part II

Inferring a Virtual Observing Agent

The arch-shaped melody is pervasive in Western music. Its common occurrence is likely the result—or perhaps the origin—of the compositional paradigm of departure and return. Theories of harmonic function posit a cyclic harmonic syntax in which a phrase departs from the stability of tonic, creates tension by traveling to the predominant and dominant, before finally resting with the return tonic at the completion of the phrase ending with an authentic cadence. This harmonic component combines with the energetic intensity associated with pitch height. As the pitch element of a melody rises, its intensity grows. This musical energy then often dissipates as the melody descends to the tonic.

Fred Lerdahl and Ray Jackendoff (1983) observe that the vast majority of works within the Western tonal idiom have a single underlying tensional shape. They describe an asymmetrical arch that has its zenith near the end. A piece typically begins with a general sense of relaxation that is prolonged. From here, tension slowly builds until reaching its highest degree. It is then immediately followed by a relatively rapid deflation, a point of relaxation that coincides with the completion of the phrase or even the work as a whole (197-98). Their observation of this musical shape that has a late climax likely comes with little surprise to the reader; however, its importance is not to be overlooked.

David Huron (1996: 3-23) has conducted an extensive analysis of melodic contours to test the accuracy of our intuitions about musical shaping; he places his attention on melodic contour. In his study, Huron draws melodies from a corpus of thousands of European folk songs and subjects them to a calculus that he has devised to determine the shapes of phrases of varying lengths. He assesses phrase size by the number of articulated pitches in each melody and maps pitch height in semitones. Huron's conclusions validate our intuitions about the prevalence of arch-melodies, since this was by far the most common shape. However, it is notable that his study does *not* reveal the melodic contour with a quick ascent and slow descent that has been our focus. The scarcity of the quickly-ascending-slowly-descending (hereafter **QASD**) contour is thus beneficial to narrowing our interpretation of it.

Markedness

A sign that has a feature setting it apart from others like it is considered to be marked, or to possess *markedness*, meaning that such signs have a particular value in comparison to their unmarked counterparts. The marked term in an opposition will typically have a narrower range of meaning. Edwin Batistella (1990) demonstrates markedness in the opposite pair *tall* and *short* (3). The opposition in these two words seems clear and unremarkable. People of great height are considered *tall* while those below average height are deemed *short*. However, in describing the height of any person, we are given to say, "She is five feet *tall*," even though five feet is below average in most populations. In a more extreme example, we might say, "She is three feet *tall*," still

invoking one term of the opposition despite the diminutive stature. This makes it clear that the descriptor *tall* refers not only to great height but that it is also used in reference to height in general; thus, *tall* is the unmarked term of the opposition. Conversely, *short* lacks the latter usage. We do not often say, “She is five feet *short*.” Such a turn of phrase adds an element of commentary that *tall* lacks in the same context. Although lexically oppositional, *short* is the marked term, and has a narrower meaning that is reserved for emphasizing a small degree of height. *Tall* on the other hand has a broader range of meaning, referring both to great height and to height in general.⁵

Frequency also plays in defining the marked/unmarked relationship. Joseph Greenberg (1966) notes that marked features generally occur less often than their unmarked counterparts. Regarding language, Greenberg writes:

[T]he number of phonemes with a particular feature is never greater than the number of phonemes with some other feature, [sic] it generally seems to be the set characterized by the marked feature which is less than or equal in number to the set with the unmarked feature.

(1966: 21)

⁵ Markedness features significantly in Hatten’s early work on meaning and interpretation (1994). He explains that the unmarked and marked elements of an opposition (e.g., major and minor) contain an asymmetry in breadth of meaning. For example, Hatten demonstrates that the minor mode is marked in the Classical style when compared to its unmarked major counterpart. The asymmetry exists because minor keys express a consistently narrow, tragic affect, while major keys are broadly characterized as non-tragic (34-38).

Linda Waugh (1982: 302) likens the marked/unmarked relationship to the figure/ground relationship in that marked elements are rarer and more forward in our perception than elements in the background. Additionally, figures are often more articulate—narrower and more precise in meaning—than the grounds against which they are placed. Schenkerian analysis reveals this in music. The *Ursatz* that underlies a phrase is generally unremarkable, since many works are built from same background structure. It would be nearly impossible to identify a piece by looking only at its fundamental structure, because only a handful of fundamental structures occur in all tonal music. The way a piece differentiates itself from others is through the distinctiveness of its foreground, the unique features that are present at its surface. Because these foreground features are novel to each piece, the surface (figure) is marked relative to the commonplace fundamental structure (ground).

The QASD (quickly-ascending-slowing-descending) gesture that I have identified above does not have an opposite in the same way that *tall* has the opposite *short*. Height exists on a clear continuum of which *tall* and *short* describe opposing sections at opposite ends of a somewhat arbitrary median. Furthermore, *tall* and *short* describe only a single feature, that of height. The QASD gesture on the other hand comprises multiple features, most notably a height component in addition to a temporal component associated with horizontal distance. A melodic arch with a slow ascent and a quick descent is not only oppositional to the QASD shape, but it is also oppositional to a U-shaped melodic contour that begins with a quick descent and slowly rises to completion. The word *girl* is analogous because it represents two features: gender and age. *Girl* is the opposite of *boy*

in the first sense while *girl* is the opposite of *woman* in the second. In the same way, the QASD gesture does not have a solitary opposition against which it is marked.

But how is the QASD contour marked if it lacks a single unmarked opposition?

On this Battistella writes:

At the semantic level of language, markedness is probably most easily understood as a relation between a very specific linguistic sign (the marked term) and a sign that is unspecified for the grammatical conceptual feature in question. In this sense, marked and unmarked elements are not strictly opposite. (1990: 2)

With regard to melodic shape, the QASD contour has a very specific and rare feature—an early apex. Huron’s survey of melodies has shown us that although most melodic shapes unfold as arches, their high points most often occur at or after the midpoint (1996). Thus, the marked QASD contour is set in opposition to these other melodic contours by possessing an earlier apex. Battistella writes that “[o]ppositions between the presence and absence of a feature—between A and not A—are referred to as *privative oppositions*” (1990: 2, emphasis mine). Logically, “A” represents the QASD contour while “not A” refers to all other contours. Appropriately, we then consider the QASD contour with its narrow range of meaning and distribution as marked against the unmarked range of meanings of the more common contours that are “not QASD.”

The markedness of the QASD melodic contour also corresponds to the markedness of the gazing gesture; they each have a similar spatial and energetic shape

and a narrower range of interpretation. If someone were to watch us collectively lift our head, neck, and eyes rapidly and then slowly lower them, the person standing before us would likely think that our attention is drawn toward something else, perhaps something that is in the distance. We recall from Chapter 1 that Marion Guck suggests that the gesture, as set in the B-minor prelude, represents labored breathing (1981: 37). Although this seemed to broaden the range of meaning of the gesture and rob it of its distinctiveness, I argued that the musical gesture could be a multivalent signifier with two complementary somatic interpretations. It is easy to imagine an agent lifting and lowering the head, neck, and eyes in accordance with the melodic contour while breathing deeply with the same energetic shape. As the agent quickly takes a deep breath, it typically directs its vision upward. The agent exhales at a more measured pace, lowering its shoulders as the head, neck, and eyes gradually descend. With this, the two interpretations of the QASD contour merge within the body to create a more visceral embodiment of the prelude's expression.

From Listener to Virtual Agent

Having explored the markedness of the QASD contour that narrows its range of meaning, we are poised to further situate it within a theory of embodied observational gesture. To accomplish this most effectively, I follow the chain of inferences that one makes when connecting a real listener to a virtual observing agent. This exercise in logical reconstruction attempts to clarify any present ambiguities in the theory. I present the inferences of the theory in two sections. In the first, I take a step-by-step look at the logical inferences while simultaneously demonstrating their application to Chopin's Prelude in B Minor. In the second section, I apply these inferences to Chopin's Barcarolle in F-sharp Major, Op. 60.

Case Study 1: Chopin, Prelude in B Minor, Op. 28, no. 6

1) *An actual agent (listener) hears a musical gesture as embodied (human) gesture.*

Arnie Cox (2001, 2011) holds that listeners hear music as mimetic of human movement through a physical empathy, whereby musical sounds are conceptualized through the listeners' own motor systems. According to this "mimetic hypothesis," in the process of listening, a listener

- a) performs an off-line simulation of the movements necessary to produce the sound, or
- b) performs an off-line simulation of a similar movement with the same energetic profile.

In the first instance, an engaged listener might hear the sound of drumming and imagine beating the drum herself. This movement may even be realized through physical entrainment. Certainly, we have all caught ourselves or others making micro-movements with the hands to imitate a prominent drum part in a piece. In the second instance, a listener who hears an upward melodic leap followed by a step-wise descent (i.e., an *appoggiatura*) may simulate (or perform) an action such as lifting and lowering the shoulders or eyebrows, or she might imagine reaching upward with the arm and grabbing something with the hand.

Chopin, Prelude in B Minor

The melodic arch in mm. 1-2 has bodily implications in that the listener hears the upward and downward motion unfold through space as a unified shape that is tied together by a phrase mark. The gesture is further unified through a dynamic arch; the natural increase and decrease in dynamics adhere to form a single gestural shape that follows the familiar tension-and-release paradigm. This movement is not only replicable in the human body. The mimetic hypothesis (Cox 2011) holds that our conceptualization of these gestures requires, at minimum, our off-line simulation of them. A number of somatic motions with this shape readily may come to mind, such as inhaling and exhaling (Guck 1981) or more generically, pushing and pulling.

2) *Musical gesture heard as embodied (human) gesture implies the presence of a virtual agent in a virtual environment.*

Physical gestures differ from musical gestures in that the gestural aspect of the former is real while that of the latter must be realized in sound to convey real qualities of an energetic shaping through time. In other words, physical gestures have a real materiality to them (i.e., the human body), while musical gestures are *musical representations/expressions* of physical movement. But just like their real physically-embodied counterparts, musical gestures require a materiality, albeit one that is virtual. Accordingly, when we identify a virtual gesture in music, we necessarily infer the presence of a virtual agent who embodies it.

We can follow the parallel further. In the same way that real agents who embody gestures exist in a real environment, virtual agents enact their virtual gestures within a virtual environment (Hatten 2004: 115), what I call a *diegetic space* or *diegesis*. Within this virtual environment, we experience the effects of *virtual forces* such as gravity, inertia, magnetism (Larson 2012), momentum, and friction (Hatten 2004: 115-16) as environmental forces. We can also imagine this diegetic space as having virtual boundaries (Hatten 2012: [4]) that influence the motion of the gesture. For example, the setting of a barcarolle in compound duple meter influences the gestures of the agent, causing them to unfold within a rocking, back-and-forth metric framework.

Chopin, Prelude in B Minor

The presence of the tension-and-release gesture in mm. 1-2 implies the presence of an agent who embodies it. Because this gesture is virtual (i.e., a musical representation of gesture) the agent who performs this gesture is similarly virtual. This virtual agent exists within a virtual environment that contains forces and boundaries that influence its movement. For instance, the *cresecendo* that accompanies the quick upward motion in m. 1 suggests the presence of an energetic impulse that resists the downward force of gravity. With this in mind, we can hear the *decrescendo* that accompanies the descent through m. 2 as yielding to this force. Here the virtual agent gives in and allows gravity to bring it to the ground. This state of rest from which the agent departs and later returns is represented by tonic, specifically B₂, as it occurs on beat one, a point of metric strength. This metric pitch location serves as a foundation for the agent—a point where it is at rest.

Inwardly-focused or outwardly-focused gestures?

As a sign, a gesture communicates information and directs the attention of the person receiving this information either toward (inward) or outward (away) from the gesturing agent. For example, a gesturer might lower her head and place her face in her hands in an expression of sadness. This is an inwardly-focused gesture in that it communicates information about the gesturer. In a different situation, if we ask someone

to show us the exit, she is likely to gesture with an outward arm and an index finger toward the door. This pointing gesture does not draw our attention toward the gesturing agent. Instead, we look outward, away from the gesturing agent, in the direction indicated by the outstretched arm.

There are, however, inwardly- and outwardly-focused elements of the gesture in both of the above situations. The gesture of sadness likely draws a causal connection to something outside the gesturing agent. If we were to come across this gesturer, we would undoubtedly speculate about the cause of this gesture. It could be the reaction to a reprimand from a teacher, news of an unfaithful partner, or the sight of a car accident. In a similar way, an outward-focused gesture can communicate information about the gesturing agent. When asked to be shown out, a very forceful pointing gesture toward the door reflects something about the gesturer, perhaps her annoyance and desire to be left alone. Because much information is contained in a single gesture, we must concede that the distinction between inwardly- and outwardly-focused gestures is fluid. It is through a gestural competency that we are best able to interpret the practical function of a gesture in a given context and according to an interpretive intention.

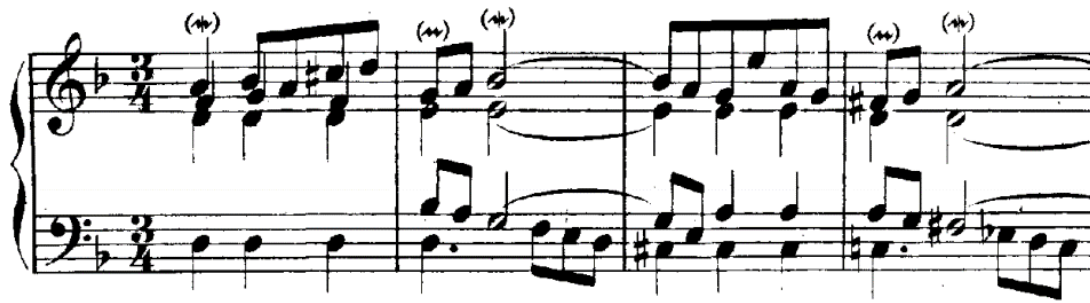
3a) *The prominence of the gestural aspect implies the presence of a virtual gesturing agent.*

3b) *The prominence of the observational aspect implies the presence of a virtual observing agent.*

It is important at this point to make a distinction between virtual agency and a virtual agent. Although we often overlook the distinction, it is revealing and profitable to analysis to describe their relationship. I define *virtual agency* as the capacity for an inferred entity within a virtual world (environment or diegesis) to act with an apparent volition and self-awareness. A *virtual agent* is this entity that possesses these inferred qualities. Examples of virtual agents include the novice in Dukas's *The Sorcerer's Apprentice*, Florestan and Eusebius in Schumann's *Davidsbündlertänze*, and the personification of Fate in Beethoven's Symphony No. 5. An agent may also go without a nominal designation, as is the case in Beethoven's *Sonata Pathétique*, Op. 13, which begins with particularly strong emotions that are expressed in both harmonic tension and articulation. Through its gestural and emotional profile, we infer the presence of a virtual agent who enacts the implied physical movements and experiences the feelings heard in the work, yet we do not ascribe a specific name to this agent. In each of these examples, agency is found in musical action and expression, and the agent results from the coalescence of these agential actions into a unified virtual entity.

Distinguishing virtual agency from the virtual agent permits us to recognize two agential modes: gesturing (or performative) agency and observing agency.

Gesturing agency describes somatic movements and emotions expressed in music that are at the center of a musical discourse. Dance genres include excellent examples of gesturing agency because they have a clear performative element as Bach's first Sarabande from his English Suite No. 1 (Example 2.1) demonstrates this aspect. The weight on beat two of each measure is a defining feature of the slow dance in triple meter. The short-long rhythm that we hear in mm. 2 and 4 gives the music the gestural weight that is reflected by a dancer whose actual weight shifts accordingly. As listeners, we center our attention on the dancer and the dancing gestures that he or she performs. In contrast, *observational agency* is present when agential gestures and emotions are directed *outward*, toward an entity that is inferred to be at a distance from the virtual agent. These entities include virtual objects, environmental boundaries or gravitational "platforms" (Hatten 2012), and other agents. Although the gestures and emotions that produce this type of agency are still embodied by the agent, they are importantly *indexical* in that they draw focus away from the agent and toward another (detached) element within the diegesis.



Example 2.1. J.S. Bach, English Suite No. 1 in D Minor, BWV 812, Sarabande, mm. 1-4

Having teased apart the distinction between an agent and its agency, we now bring the two together. Gestural agency and observing agency are agential modes that are present in two distinct *agential roles*: the virtual gesturing agent and the virtual observing agent. The *virtual gesturing agent* is the role taken on by the virtual agent when it exhibits gestural agency, such as by performing gestures with an immediate focus to itself. In hearing the dancing gestures in the rhythm and meter of the sarabande, we imply that an agent—a dancer—also exists to embody the movement. The *virtual observing agent* refers to the virtual agent's role when performing indexical gestures such as pointing or looking. The identification of agential roles is important because it permits an agent to both *act in* the diegesis (e.g., run, jump, and dance) while being *observant of* entities in the diegesis (e.g., look and point).

Chopin, Prelude in B Minor

In her work on this prelude, Guck (1981) suggests a number of guiding metaphors for listening. In each, she suggests a guiding metaphor and then allows her subjects to further construct an interpretation. What she reveals in these studies is that listeners can fit music to a number of interpretive metaphors, each drawing attention to different salient features of the musical structure. Guck eschews the notion that with any piece there is a “correct” metaphor, writing

As patterns flow together, they combine their power, inexorably drawing the listener’s attention. Compelling as they are, such confluences inevitably reach a peak of interaction before each flows off at its own rate. As patterns flow each in their own way, they may work at cross purposes to each other, eddying and conflicting. *Patterns rival each other for the listener’s attention*, and the listener senses the different currents as they surge to the surface and then submerge, flowing against or through each other and parting (1981: 123, italics my own).

With this, she suggests that an interpretation may benefit by mixing metaphors and using the strengths of each to help us better conceptualize the piece.

- 4) *The virtual observing agent individuates itself from virtual objects, a virtual environment, and other virtual agents.*

The indexical gestures of the virtual observing agent imply a distinction between itself and other entities that exist with the diegetic space of the music. After all, the act of observation requires two essential elements: the *observer* and the *observed entity*. Consider the following scenario:

When I first moved to Austin, Texas for graduate study, I learned my way around the city by jogging through a different neighborhood each day at dusk. On my first evening run through downtown, I noticed several hundred people lining the east side of the city's most prominent bridge. From a distance, I assumed that these were demonstrators whose protest zone had been relegated to one side of the bridge so as not to impede pedestrian traffic. However, as I approached the multitude, I noticed that everyone was facing outward from the bridge with their eyes directed slightly upward. Their direction of sight and pointing gestures led me to conclude that they were *observers*. These gestures drew my attention away from the crowd and toward something that was apart from it. As I followed their gaze, the *observed entity* was revealed to me: thousands of Mexican free-tailed bats flying in a vortex pattern above the trees. It was a remarkable sight.

The gestures of the virtual observing agent cue in the listener to the details of the entities that it observes. Once I had understood that the people on the bridge were performing observing gestures, I was able to deduce a number of things about

what they were looking at *before* actually observing it myself. The upward gaze and the moving heads of the bat-watchers led me provisionally to assume that objects that they were observing were airborne and still in motion. I further deduced from the emphatic pointing gestures of the children that the object of observation was something remarkable; they were probably observing something that to this point they had never seen.

Observing gestures can also reveal information about the environment. Mentioned above, Hatten (2012) characterizes the musical forces described by Larson (2012) as *environmental forces*. Larson convincingly demonstrates the way in which a melody seems to act like a physical object that is subjected to the forces of gravity, magnetism, and inertia. According to this model, melodic pitches are simultaneously drawn downward (gravity), toward a point of stability (magnetism), and at the speed and in the direction of on-going motion (inertia). Matthew BaileyShea, however, suggests that forces specific to a particular environment can also influence musical motion (2012: [11]). For example, the compound meter of barcarolles and gondola songs reflects the rocking of the boat that is central to those genres.

Chopin, Prelude in B Minor

The existence of a gesture assumes the presence of an agent to embody it. The act of gazing, for instance, requires an agential entity to perform the gaze. Similarly,

pointing cannot occur without implying the existence of an entity that enacts the pointing. Thus, gesture implies agency.⁶

5) *The individuation of the virtual observing agent implies a single observational perspective of the diegesis.*

In embodying the individuated virtual observing agent, the listener is given a perspective of the diegetic space. The listener constructs the virtual world represented in the music as though he or she were positioned uniquely within it. The geometry of the imagined virtual space appears skewed from this perspective. Objects increase in size and appear more detailed as they come into the foreground; parallel lines converge as they extend into the distance. This relational space is created in various ways. Differences in dynamic intensity can distinguish objects that are close from those that are distant. This is true of music, as well. For example, the (less intense) vamp of a rhythm section that begins a piece of jazz music provides a setting that is generally unremarkable when compared to the (more intense) highly configured theme that emerges. If we subject the passage to a Schenkerian analysis, we see that more mundane parts of

⁶ Although gesture implies agency, this agency may be subtle. For instance, an electronic traffic sign with a flashing right-pointing arrow in a construction zone is without a doubt an object; however, agency is embedded within it because the sign "tells" the driver to merge right. Although it is the arrow (i.e., an object) that is instructing us, we subtly ascribe agency to it, understanding it as a command from an authoritative agent.

the music are placed in the background while the attention-grabbing, unique elements comprise the foreground.

Chopin, Prelude in B Minor

The quick-ascent-slow-descent (QASD) shape of the melodic arch in mm. 1-2 resembles the bodily movements that one would make when tracing the path of an object moving into the distance. As demonstrated in the first part of this chapter, the *actual* flight path of such an object is parabolic. Accordingly, one who is positioned directly beneath the apex of motion, or at a point that is perpendicular to it, tracks its motion with a symmetrical gesture resembling the parabola. However, it is the unique shape of the gesture in these measures that positions that virtual observing agent at a specific point relative to the object's motion. In drawing a connection between the QASD gesture and the collective motions of the observer's head, neck, and eyes, we come to understand the perspective from which this virtual agent views the virtual object.

6) *The individuation of the virtual observing agent implies a single subjectivity possessing a single emotional perspective or unique affective state.*

By individuating the virtual observing agent from other entities within the diegesis and by locating this agent in relation to those entities that it observes, we in effect position a lens through which a listener observes this diegetic space. It

might be helpful to conceptualize the virtual observing agent as a coin-operated tower viewer that one would find atop the Empire State Building. These stationary telescopes offer a view of the urban landscape from a solitary position. In the same way that a tourist views the city below through the mounted binoculars, the real listener observes the diegetic space of the work through the virtual eyes of a positioned virtual observing agent.

Like most metaphors, the above one has its limits. Whereas the tower viewer only contains lenses to give the viewer visual information, the human-like virtual observing agent can transmit all types of sensory information to the real listener. For example, dynamics can portray relative physical distance. Mahler uses this technique in Symphony No. 1 when he sets the trumpets offstage at the beginning of the first movement.⁷ By positioning the real trumpets at a distance from the audience and reducing their volume, he places the *virtual* musicians that perform the heralding call at a distance, within the diegetic space of the pastoral scene. Meter and rhythm can similarly reflect the bodily movement of a virtual agent. An example of this is the oboe recitative from the Bacchanale of Saint-Saëns's *Samson and Delilah* that captures the sensuousness of the orgiastic genre (Example. 2.2). Free from the rigidity of meter (*ad lib.*), one can hear the sinuous melody as depicting the unsteady body of a virtual agent caught up in drunken

⁷ The score indicates that the trumpets should be “set at a very far distance” (*In sehr weiter Entfernung aufgestellt*).

revelry. The melody begins with vain attempts to stabilize itself on A₅, slipping numerous times in the process. The line then slurs downward and unsuccessfully grabs at E₅ before falling finally to A₃ where it comes to rest. Alternatively, one might just as easily hear the exotic Phrygian cadenza as expressing a sultry, sexual bodily movement. Regardless of the interpretation, it does not take much effort for an attentive listener to imagine himself or herself as performing these actions. By experiencing the music as through one were a virtual agent (or virtual observing agent), the real listener imagines the complex of sensations (e.g., vision, sound, movement, balance, etc.) of that virtual agent.



Example 2.2. Saint-Saëns, *Samson and Delilah*, Bacchanale, recitative

The collection of these sensations within the virtual observing agent suggests various thoughts and emotions—hence, an emerging subjectivity. This subjectivity within the virtual *observing* agent is unique in that it is positioned within the diegetic space. It observes only what is made observable; it is not omniscient. The affective state is often similarly positioned. The virtual observing agent usually experiences emotions that are aligned, because these affective feelings reflect its unified and individuated subjectivity.

Chopin, Prelude in B Minor

The unique melodic contour of the left hand in mm. 1-2 uniquely positions the virtual observing agent within the diegetic space of the work. As the virtual object travels away from the virtual observing agent, the agent experiences a feeling of loss or deprivation. The articulation of this affect is enhanced by the “tolling bells” or consoling hand gestures found in the highest voice and, of course, the slow tempo and minor mode. The attentive listener follows the entailments of the agent’s gaze within the topical environment to make better cognitive and emotional sense of the agent’s subjectivity—its conscious self. With this final step, we see that the music implies the presence of an individuated, sensing, thinking, feeling, and conscious, human-like agency, through whose skin the listener can experience the music.

Case Study 2: Chopin, Barcarolle in F-sharp Major, Op. 60

Chopin’s Barcarolle is another pensive (though less gloomy) work for solo piano that features a QASD melodic shape. Like the Prelude in B Minor, the Barcarolle projects a wistful and longing character. The nine minutes that pass between the first measure and the last seem to represent a space that is almost devoid of time. The 12/8 meter offers the piece a delicate sway that is filled in by almost constant eighth-note rhythms. This

permits both the expression of an observer gazing at the night sky, and the buoyant rocking of the gondola in which we might imagine the observer sitting.⁸

The piece opens with low C#s that are sounded in the first and second octaves. Their earthy tone resonates deep in the listener’s chest, providing a solid foundation and an embodied point of reference for what is to come. The next beat presents a stark contrast to the first with G#₅. The melody maintains its height for a moment then descends lazily into m. 3. Its indirect fall is like that of a feather; the melody flutters upward and downward while making its protracted descent. The inner voices complement this light and ethereal sound by creating consonant intervals on each of the following beats. An appoggiatura on each beat injects a soft energy that pushes the music forward.

The image displays a musical score for the first six measures of Chopin's Barcarolle in F-sharp Major, Op. 60. The tempo is marked 'Allegretto'. The score is in 6/8 time and F# major. The first measure begins with a forte (f) dynamic. The second measure features a decrescendo (dim.) dynamic. The third measure is marked 'cantabile'. The bass line consists of eighth notes with appoggiaturas, marked with piano (p) and asterisks. The treble line has a melodic line with a long note in measure 3.

Example 2.3. Chopin, Barcarolle in F-sharp Major, Op. 60, mm. 1-6

⁸ Janice Dickensheets observes that “the Venetian Gondolier song with its lilting six-eighth [sic] meter and soulful tune” conjure thoughts of “a simpler way of life, neither too rural nor too exotic” (2012: 131).

David Kopp (2014) describes the three measures that begin this piece as depicting a pyrotechnic device—a rocket—as it soars into the night sky above the canals the Venice. His vivid account is as follows:

The barcarolle as a genre is, of course, a Venetian boat song. One of my most lasting impressions of a long-ago trip to Venice is of a sultry summer night which happened to be the festival day. Over the canals, fireworks repeatedly lit up the sky. Not until years later did this image occur to me as I played the opening measures of the Barcarolle: shot from below, the music lands high in the sky in a burst of predominant energy, its sparkling parts slowly floating down through the air at slightly unequal speeds. [9]

Once heard in this way, it is difficult to imagine a more memorable and appropriate scene. The octave C-sharps on beat one of m. 1 portray the initial combusive force that propels the rocket upward. The depth of the sound implies both the thud of the rocket as the initial propulsive force is applied to it and it leaves the mortar barrel. In the silent space between beats one and two, we visually track the almost visually imperceptible ascent of the pyrotechnic. The right-hand entrance of the supertonic G# minor 6/3 chord on beat two portrays the mid-air explosion, causing a sudden burst of light that blooms in the night sky. The slow-moving descent into m. 3 vividly illustrates the pockets of light

casually falling to the earth, with the pitches of the harmonic texture sketching the many glowing embers in the mind of the listener. I have rendered Kopp's description in Figure 2.8.

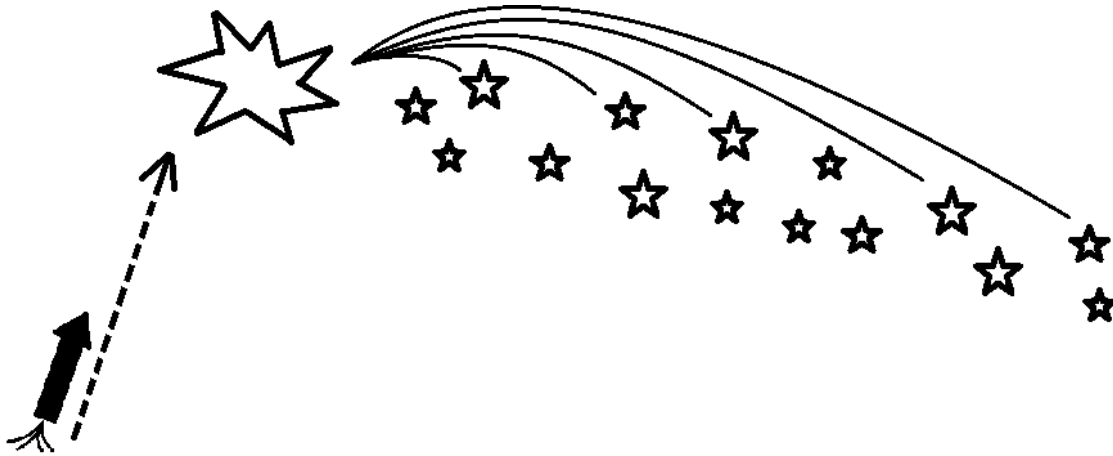


Figure 2.8. Visual representation of Kopp's description of Chopin, Barcarolle, mm. 1-3

The pyrotechnic rocket that we hear in the first three measure of his piece is a *virtual object* since it is the object toward which the virtual observing agent directs its gaze. Unlike the virtual object depicted in the Prelude in B Minor, the virtual object in the barcarolle has *virtual materiality within the diegesis*. It is not simply a metaphor for a past memory or an uncertain future; instead, it exists in the same virtual space that contains the virtual observing agent. It is already a rocket at its first level of interpretation.

In a few sentences, Kopp is able to describe the three-measure introduction in great detail. He presents a vivid image that is immediately accessible to the listener, even one who has not experienced fireworks over the Venetian canals first hand. Kopp does

not nuance his interpretation; instead, he allows us to fill in the gaps, making connections to the music and discover the finer details ourselves, as I have attempted above. Although connecting the few dots that Kopp provides does present a challenge, it requires us to follow the same chain of inferences necessary for our interpretation of the Prelude in B Minor.

1) *Actual agent (listener) hears musical gesture as embodied (human) gesture.*

Kopp describes the Venetian scene with articulate imagery for which he hears a parallel in the Barcarolle. That Kopp hears musical sounds as representing embodied movement should come as no surprise. Recalling the mimetic hypothesis, we know that when a listener hears music, he or she conceptualizes it through either on-line or off-line motor action. We might hear the rowing motion of the gondolier that begins in the left hand in m. 4. The rhythmic and melodic contour reflects effortful push of the oar and its relaxing return to the gondolier. It is clear from these interpretations that we are hearing musical motion as physical human gesture.

2) *Musical gesture heard as embodied (human) gesture implies the presence of a virtual agent in a virtual environment.*

Although Kopp was physically present when he witnessed the fireworks soaring above the canals that one summer evening, the movements or gestures that he makes while reliving this experience musically are not necessarily physical.

While listening, he may recreate some of the physical gestures (e.g., gazing) that he performed when he was actually in Venice; however, many of these movements will be muted if not entirely inhibited. Although Kopp is a real listener, his interpretation reflects the actions and experiences of a virtual agent in a virtual environment, namely Kopp's past self that very summer evening in Venice.

3b) *The prominence of the observational role implies the presence of a virtual observing agent.*

There is a point that has yet to be made clear, one that is slightly muddled by the language that Kopp chooses to describe the Barcarolle. In writing that "the music lands high in the sky," it may seem that Kopp suggests that the music is a direct representation of the aerial firework. However, we must remember that in describing the Venetian scene, Kopp places himself in a clearly observational role. His memory is not of being the firework; it is of witnessing it. Because his

experience of its movements is mediated through his observing gestures, the agent that he posits in his description has a notably observational role.

4) *The virtual observing agent individuates itself from virtual objects, a virtual environment, and other virtual agents.*

Kopp establishes himself at a remove from the motion of the fireworks in the scene with the language that he selects to describe it. He keeps his distance by writing that “*the music* lands high in the sky” [italics mine], notably refraining from say, “I land high in the sky.” In describing the represented object (i.e., the firework) in the third-person, he individuates himself from it and establishes himself as an *observer* to its motion.

In the above analysis of the Chopin’s Prelude in B Minor, I suggested that the quick-ascent-slow-descent (QASD) melodic contour represented the physical actions necessary to track an object moving away from an observer and into the distance. Having established this as an indexical gazing gesture, I proceeded to integrate this with the solemnity of the piece and its funereal signs in order to articulate a reading that was otherwise unattainable. I concluded that the piece projects a feeling of loss through the sorrowful, longing gaze of the virtual observing agent as it tracks an unknown object moving outward from it. The

analysis sought to better define the emotional state of the virtual observing agent than was possible through purely topical interpretations.

In our focus on the virtual observing agent, I neglected to specify the virtual object—the thing that the virtual observing agent tracked with the collected movements of its head, neck, and eyes. I speculated that this virtual object could, itself, be a symbol interpretable as a metaphor for a distant memory or a vision of the future. Regardless of the interpretation, I recognized the implied contemplation bound up in the gazing gesture. However, I purposefully avoided further describing the virtual object as a material entity within the diegetic space because the music lacked the necessary degree of specificity for me to unequivocally identify and name it. (Its characterization as a ball was hypothetical and served only to aid the understanding of ballistic motion.) This ambiguity drew my attention to the virtual object’s symbolic meaning. I concluded that the QASD contour represented a longing gaze through time (toward the past or future), but I stopped short of further describing the virtual object, holding that the music articulates its symbolism to a greater extent than its virtual materiality.

5) *The individuation of the virtual observing agent implies a single observational perspective of the diegesis.*

The position of the virtual observing agent relative to the virtual object (i.e., the firework) is not overly specified by the music. We cannot say that the virtual

observing agent sits in a gondola at the Rialto Bridge and that the firework is launched from a position 510 feet north and blossoms in the sky at a height of 212 feet. This music simply does not provide this level of detail in its construction of the diegetic space. Yet the music is not so vague that it does not provide some important information about the virtual environment. We can deduce that the firework travels along an upward path that draws the observer's vision upward. The speed and degree at which the gesture occurs suggests that the virtual observing agent and the virtual object are relatively close to one another.

6) *The individuation of the virtual observing agent implies a single subjectivity possessing a single emotional perspective or unique affective state.*

The observing gesture in mm. 1-3 positions the virtual observing agent at a point below the virtual object—the aerial firework. When informed by the topical allusion to the gondola and motivated by Kopp's recollection of an evening in Venice, the unique contour of the QASD gesture provides us with clues to the nature of the object and perspective from which it is observed by the virtual agent. Moreover, the agent also experiences a complex of emotions, thoughts, and other sensations while in this setting. This includes the rocking of the 12/8 meter, particularly palpable in the accompaniment that begins in m. 4. Through the virtual observing agent, the listener may feel pensive or wistful both through the far-off gaze of the QASD gesture and by the simplicity suggested by the genre

(Dickensheets 2012: 131). Certainly, an analysis of these few measures falls well short of offering a thorough interpretation of the work. But from this small section of music, it is possible to demonstrate a method of conceptualizing this piece and hearing it through the body and senses of a virtual observing agent.

Conclusion

My examination of the QASD gesture has contributed to the theory of the virtual observing agent by grounding an imaginative reconstruction of motion with physics and mathematical articulation. I have demonstrated that this melodic contour imitates the physical indexical movements that an agent performs while visually tracking the path of an object that travels upward and outward. This human gesture is also found in *metaphorical gazing*, whereby a person has a far-off look while deep in thought. In many cases, both the literal and metaphorical virtual gazes are mutually supportive, reinforcing the same interpretation of the physical gestures.

CHAPTER 3

OTHER INDEXICAL GESTURES AND PERSPECTIVES

Introduction

In this chapter, I further explore musical gestures that suggest the presence of a virtual observing agent. I first place my focus on ascending melodic lines that mimic the indexical gestures of an agent casting an upward gaze or reaching outward with aspiration or desire. I find that in certain musical contexts, this indexical gesture toward a virtual object can lead the listener to identify an emotional yearning, a type of gazing in which the “object” is metaphorical within the diegesis. In my examination of these musical gestures, I also consider the performance gestures that they reify, notably with reference to Beethoven’s *Emperor Concerto* and his Piano Sonata in A Major, Op. 101. Further on in this chapter, I shift the relationship between the observing agent and the virtual object, studying examples in which the virtual object remains in a fixed position while the observing agent moves around it. I propose a virtual observing agent that takes in the sight of a stationary object from multiple perspectives. Through this interpretive lens, I look at two compositions programmatically linked to monumental architectural structures: Schumann’s Symphony No. 3, iv, and Debussy’s *La cathédrale engloutie*.

Part I

Yearning as Indexical Gesture

Indexical Gesture in Beethoven, Piano Concerto No. 5 (*Emperor Concerto*), Op. 72, ii

The second movement of Beethoven's *Emperor Concerto* is one of the most beautiful pieces of the composer's catalog. Nearly fifty years after its first performance, biographer Anton Schindler considered the work as a whole to be the "summit of all concerto music written for the instrument," citing the composer's skill in projecting spirituality (1966 [1840, 1860]: 160). William Kinderman observes that "the serene slow movement in B major acts as an immense parenthesis," describing its mood as "dream-like reflection" (1995: 135). In Example 3.1, we see that the movement contains a number of topical markers associated with the expression of pastoral bliss: slow tempo (*Adagio un poco mosso*), major mode, soft dynamics, parallel thirds (violins, mm 1-2, 9, 11), contrapuntal wedges (mm. 7, 9, 11) and consonant appoggiaturas (mm. 5-6, 13-15) (Hatten 1994: 98-99). The movement is non-virtuosic, as typical of concerto slow movements. Its demands from the pianist instead lie in the delicacy that one must employ while navigating the keyboard. The melody further reflects the simplicity of the idyll (97). It is through the careful attention to timing and the recognition of the muscular nuance required of each note that the pianist is able to fully realize the expression of the piece.

Adagio un poco mosso.

Violin I
Violin II
Viola *p*
Bass

5 Woodwinds
cresc. *f*

10 *dim. p* *f* *p* *p*
V⁷ vi V⁷ I

16 Piano
pp *espressivo*
Strings

Example 3.1. Beethoven, *Emperor Concerto*, Op. 73, ii, mm. 1-17

The movement begins with the homorhythmic texture and controlled counterpoint of the bound style. In m. 5, the melody of the first violin jumps upward a minor seventh from F#₄ to E₅. Following the chordal texture of the first four measures, the sudden assertion of melodic independence by the highest voice catches the listener's attention. The energy heard in this quick and substantial melodic leap in the first half of m. 5 seems to dissipate in the second half of the measure as the melody descends more leisurely E₄ – D#₄ – B₃ – F#₃. This gesture repeats in the following bar (m. 6). The importance of this repetition and reinforcement become understood in m. 16 where the solo piano has its first entrance. Here, the melody also makes a rapid ascent from F# in the fifth octave, this time, to F# high in the sixth octave. The short duration of the lower F# seems to reflect a sort of propulsion, much like the firework that David Kopp (2014: [9]) uses to describe the opening gesture in Chopin's Barcarolle, Op. 60. It hangs for a moment and descends leisurely with triplets in a simple meter. The metric discontinuity (three-against-two) allows the melody to float downward, rather than ascend more directly. The rhythmic non-alignment of the latter two notes of each triplet figure as though the melody were a feather in mm. 16-17 offers it a degree of lightness even catching an occasional updraft while floating downward.

Although the imagery of a feather or a firework is easy to visualize, it remains detached from the human body and, therefore, it does not directly represent a human gesture. We recall that in Chapters 1 and 2, this quickly-ascending-slowly-descending (QASD) melodic contour reflected the collective head, neck, and eye movements that an

observer makes when visually tracking an object that moves into the distance along a parabolic path. The visual tracking movements are much quicker as the object ascends in the observer's foreground of vision, but more measured as the object descends into the background. Interestingly, this is the gesture that one might make when suddenly encountering a beautiful vista. After all, pastoral depictions in art have long represented an idealized vision of nature while providing a setting for deep thought and self-reflection.¹

We can imagine the virtual agent coming upon a scene of verdant pastures that stretch outward toward the horizon. Faced with the impressive sight, the subject's head darts upward quickly, perhaps accompanied by a sudden awed gasp, before descending at a slower pace. The gesture also is one that might occur if one is seized by a sudden insight and then reflects upon it in wonder, as in metaphorical gazing.

Amid the contemplation of this melodic arch, we find another indexical gesture in the concerto, one that is more pressing in character. In mm. 7-9, the first violin pushes upward from the weight of F#₄, through G₄, to G#₄ ($\hat{5} - \hat{b6} - \hat{6}$). The tonal stability of F# ($\hat{5}$) gives it a magnetism that resists the upward motion to G and G# (Larson 2012: 88-95).² The laborious chromatic ascent and the *crescendo* are evidence of this effort.

¹ Renato Poggioli notes that European writers of the seventeenth century shifted the depiction of the pastoral from one of outward passion to one of inward contemplation and self-examination. By exploring "pastorals of melancholy and solitude" these authors were able to develop "the pastoral of the self, which in the end transcended all previous traditions of the genre" (1959: 686).

² Because the tonic triad creates tonal stability, Larson identifies its component pitches as exerting a metaphorical magnetism. He writes that "[m]elodic magnetism is the tendency of an unstable note to move to the closest stable pitch" (2012: 88).

After pausing to gather strength, the melody strives upward again, reaching its tonic goal (B), on beat 3 of m. 8. These linear ascending gestures are much different from the relaxed gaze of mm. 5-6 and 16-17. The melody in mm. 7-8 expresses the effort of a virtual observing agent as it *reaches* or *yearns* for tonic.³ In embodying this agent, one stretches the arms outward toward something that is beyond reach. The virtual observing agent indexes the object of desire by “pointing” its arms and focusing its attention in that direction.

Once reaching tonal stability, the melody climbs upward yet again; this time the woodwinds double the violin’s stepwise ascent $B_4 - C\#_5 - D\#_5$ ($\hat{1} - \hat{2} - \hat{3}$). This gesture expresses a reach toward spiritual fulfillment that goes beyond what is tonally sufficient (B) in order to achieve a sense of transcendence (D#). This spiritual achievement, however, proves elusive, since the deceptive cadence ($V^7 - vi$) in m. 10 undercuts the expected perfect authentic cadence. The virtual agent then makes another attempt toward cadential fulfillment and achieves it with $V^7 - I$ across mm. 12-13. The suddenly soft dynamics of those measures, however, rob the music of the spiritual perfection that it seeks; of course, such an achievement is not to be expected in the opening bars of a lengthy movement. Instead, the early placement of this gesture that yearns for spiritual

³ In her article “Sources, Containers, and Paths,” Janna Saslaw (1996: 217-43) discusses musical direction using the “source-path-goal” image schema. Saslaw recounts the many ways that theorists historically have conceptualized forward harmonic, for example, “starting with the tonic and passing through a small or large number of intermediate locations, that is, chords” before reaching tonic once again (222). The yearning that we hear in mm. 7-8 draws on this same schema and identifies the listener’s foreknowledge of the melodic and harmonic goals that lie (hopefully) at the end of the musical “path.”

fulfillment permits it to serve as a premise for the aspiration expressed throughout the remainder of the movement.

We take away from this example (mm. 8-9) the notion that an ascending melodic line can project an indexical gesture, particularly when it is goal-directed. This gesture is similar to the observing gesture heard in the QASD contour (m. 7-9), but here the step-wise gesture expresses a more pressing desire as it draws nearer to the upper tonic. The gesture meets Eco's (1976: 119) criteria for indexical pointing; like the pointing finger, the melody is 1) elongated (longitude) by its 2) direction of motion, and its 3) apical tip is present in each successive pitch. The 4) dynamic stress of the gesture is present in the intensity of its projected desire. Although the gesture seems to achieve local goals (i.e., $\hat{1}$ in m. 8 and $\hat{3}$ in m. 9), several musical undercuttings keep it detached from its overall expressive goal in these measures. This reaching gesture continues to point toward the virtual object of its desire but it never quite attains this goal.

Teleological Melodic Ascent as Indexical Gesture

The step-wise ascent that progresses from $\hat{5}$ to $\hat{8}$ toward an authentic cadence is notable for its subversion of contrapuntal norms. According to Steve Larson's theory of melodic forces (2012), a melodic ascent from dominant to tonic contradicts the prevailing magnetic forces of gravity and magnetism. In Example 3.1, $\hat{5}$ in m. 7 is tonally stable within the chordal foreground (tonic) and tonal background (B Major). Its subsequent

movement upward to $b\hat{6}$ (or $\#5$) demonstrates an exertion of agency, an action against both the downward pull of gravity and the magnetic pull of stability. Its continued upward movement demonstrates the melody's energy. The melody does not possess a restful quality even when it reaches tonic at the perfect authentic cadence of m. 8. Instead, it draws our attention forward in time to the reposeful closure that we receive later in the piece.

In a lecture at the University of Texas Music Theory Symposium (April 2014), Michael Buchler delivered a paper entitled, "What Goes Up," an examination of Tin Pan Alley and classical Broadway songs that conclude with a structural linear ascent through the upper tetrachord ($\hat{5} - \hat{6} - \hat{7} - \hat{8}$). Buchler's citations include such notable tunes as "Supercalifragilisticexpialidocious" (*Mary Poppins*), "When You Wish Upon A Star" (*Pinocchio*), and "People Will Say We're in Love" (*Oklahoma!*). Each of these show tunes concludes with an ascent to tonic that suggests the infusion of agential energy, a resistance to the environmental force of gravity. Hatten (2012: [5]) describes this as "initiatory energy" and holds that its opposition to environmental forces leads the listener to infer agency. But Buchler goes on to draw out the significance of the ascending tetrachord. He contends that each song's closing ascent is part of the background structure that comprises the fundamental line or *Urlinie*. Of course, this ascent exists as a notable contradiction to the requisite descending *Urlinie* of Schenkerian orthodoxy. Buchler, however, shows quite convincingly that in this more recent style, the ascending *Urlinie* occurs with relative frequency, citing the examples above and a host of others.

Yet this ascending structure is rare within the larger context of the Western tonal canon from which Tin Pan Alley and Broadway emerge. Thus, the ascending *Urlinie* has become not only a novel way of bringing tonal and structural closure to music, it is also a marker of expressive potential.

An example of the expressivity of the structural ascent is found in the British World War II song “The White Cliffs of Dover” (Walter Kent and Nat Burton). The piece, written during Britain’s early years in the war (1941), became popular in a recorded performance sung by Vera Lynn that was released the following year. A hope for peace and liberty from Axis powers is expressed in the first verses:

There’ll be bluebirds over,
The white cliffs of Dover,
Tomorrow just you wait and see.

There’ll be love and laughter,
And peace ever after,
When the world is free.

The imagery is well-established as the opening melodic figure ascends $\hat{3} - \hat{5} - \hat{8}$ with the lyrics “There’ll be bluebirds.” This blossoming motive mimics the fairly rapid upward movement of an observer whose head, neck, and eyes rise to capture the monumental chalky bluffs of the title. This ascending gesture becomes thematic throughout the work, and notably returns in the last line of the song. Example 3.1 depicts the vocal melody in

the final bar. The gesture continues past the initial melodic ascent with a delayed melodic *descent* that is only initiated by word “over.” The observer’s gaze thus follows a QASD arch-shaped pattern that looks both up and down the massive cliffs of the song’s title.



Example 3.2. Kent and Burton, “The White Cliffs of Dover,” (1942 recording)
(transcription mine)

The final melodic line of the piece is perhaps the most noteworthy element of the structure in light of Buchler’s work. Beginning from in the antepenultimate measure, Lynn proceeds downward $\hat{4} - \hat{3} - \hat{2}$, but balks at the expected descent to tonic. She instead leaps up a fifth ($Bb_3 - F_4$), completing the line $\hat{6} - \hat{7} - \hat{8}$.⁴

Following Buchler, I hold that the ascending tetrachord is not only a structural possibility for the fundamental line, but it can be a *marked expressive indexical gesture* that mimics the head, neck, and eye motions of gazing at or yearning for a virtual object. This song ends by looking forward in time, employing *tomorrow* as the virtual object that the singer yearns for, with an almost naïve hope. We can imagine Lynn looking upward,

⁴ The same melodic contour (QASD) begins “Over the Rainbow” from *The Wizard of Oz* (Arlen and Harburg). Interestingly, Judy Garland’s performance of the ballad also concludes with an ascent to tonic, this one $\hat{4} - \hat{5} - \hat{6} - \hat{7} - \hat{8}$, as she lifts her head, gazes upward, and asks, “Why, oh, why can’t I?” Buchler (2014) cited the example in a handout.

perhaps even extending and reaching an arm toward the peaceful future for which she pines, and her gesture constitutes a metaphorical gaze.

Buchler's study of the ascending tetrachord offers us an entry point into understanding this melodic line as a musical gesture with virtual physical implications. His choice of repertoire benefits from its pairing of melody and lyrics, which together are able to carry a specificity of meaning not found in purely instrumental music. However, what he uncovers is still valuable when the text is stripped away. We are now positioned to explore Buchler's hunches about the expressive implications of the ascending tetrachord in a study of untexted works that suggest other types of indexical gestures.

The descent of the *Urlinie* (fundamental line) is taken as a natural principle of tonal music in much the same way that gravity is taken as a natural property of the physical environment in which we live. In fact, Schenkerian theorists Allen Cadwallader and David Gagné (1998) note that Schenker found the fundamental descent to be rooted in Nature:

[T]he fundamental line *descends* from the initial tone to the final $\hat{1}$. The reason for this aspect of Schenker's theory is that he considered the first tone of the *Urlinie*—the *primary tone*—as an “image” or representation of the overtone in an overtone series of which the tonic is the fundamental. He therefore considered the descending motion to be the most direct path for this “overtone” to return to its fundamental, a

direction that also corresponds to a release of musical tension. (108)

Accordingly, a three-line descent in C major, for example, descends E - D - C ($\hat{3} - \hat{2} - \hat{1}$) because this upper-voice structure is trying to return to the fundamental. (It might be more accurate, however, to say that the *Urlinie* attempts to come as close to its fundamental as possible because this upper-voice C goal is itself an overtone of the lower C found in the bass of the complete *Ursatz*.) Thus, Schenker suggests that Western classical music is born out of a contrapuntal framework that is not based simply on style, but instead rooted in a natural acoustic phenomenon.

The acceptance of this tenet of Schenkerian theory is notably challenged by David Neumeyer in his article, “The Rising *Urlinie*” (1987). In it, Neumeyer suggests that the fundamental line may ascend to tonic ($\hat{5} - \hat{6} - \hat{7} - \hat{8}$) with the same ease that it descends in Schenker’s conceptualization of the tonal structure. Through Viktor Zuckerkandl (1956), Neumeyer identifies $\hat{5}$ as the point of most “dynamic” tension; that is, $\hat{5}$ strives to reach tonic more than do other scale degrees because of its position in the middle of the scale (Neumeyer 1987: 280). Zuckerkandl describes the dominant scale degree as figuratively perched on the edge of a knife (1956: 97) and thus bound up with potential energy that it desires to expel. If tipped in one direction, the *Urlinie* will fall to $\hat{4}$ and onward *down* to tonic ($\hat{5} - \hat{4} - \hat{3} - \hat{2} - \hat{1}$). If tipped in the other direction, it will progress *upward* until it reaches the upper tonic ($\hat{5} - \hat{6} - \hat{7} - \hat{8}$).

Neumeyer explicates with the first movement of Beethoven's Piano Sonata in A Major, Op. 101. The expression of the movement is somewhat inward, marked by the performance description *mit innigsten Empfindung* (literally, "with innermost feeling"). There are also repeated undercuttings of cadential closure, inward wedging textures, and pregnant pauses. Given this inward character, the piece seems to be an unlikely candidate for an atypical structural ascent that stretches the background upper line, pulling the *away* from the bass. However, Neumeyer finds the ascending tetrachord to be the fundamental upper line (*Urlinie*) of the piece (300-01). A section of his graph is found in Figure 3.1.



Figure 3.1. Neumeyer's (1987: 300) middleground reduction of Beethoven, Op. 101, i (truncated, mm. 59-77)

Carl Schachter critiques Neumeyer's notion of the ascending *Urlinie* in the appendix of his article, "Schoenberg's Hat and Lewis Carroll's Trousers: Upward and Downward Motion in Musical Space" (1996). He claims that $\hat{6}$ has a natural tendency to fall toward $\hat{5}$, thus inhibiting the domino-effect described by Zuckerkandl (1956: 338-39).

Schachter draws on the following diagram from Zuckerkandl to press his point (Figure 3.2):

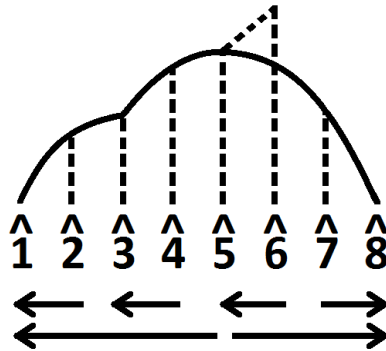


Figure 3.2. Reproduction of Zuckerkandl’s melodic tendency diagram (1956: 98)

Notice that the longer arrows on the bottom depict the “knife-edge balance” of $\hat{5}$; as mentioned earlier, it is pulled both downward to the lower tonic ($\hat{1}$) and upward to the higher tonic ($\hat{8}$). The shorter arrows just above depict the direction of force that pulls certain pitches toward more stable ones—this is similar to Larson’s force of magnetism (2012). Because pitches of the tonic triad are considered most stable, the remaining pitches of the scale draw toward them (recalling Larson’s magnetism). In considering lateral distance and the compounding influences of the arrows in the above diagram, $\hat{2}$ is drawn to $\hat{1}$, $\hat{4}$ is drawn to $\hat{3}$, and $\hat{7}$ is drawn to $\hat{8}$. The tendency of the submediant ($\hat{6}$), however, is somewhat problematic, as we have seen in Beethoven’s *Emperor Concerto*, because of the competing tendencies depicted in the opposing arrows. It is here that Schachter centers his argument, claiming that the downward pull of the top arrow

counteracts the upward pull of the arrow beneath it.⁵ Therefore, when $\hat{5}$ is “tipped” in the direction of $\hat{6}$, it returns to $\hat{5}$ in a neighbor motion. The downward pull to $\hat{5}$ overcomes the upward pull toward $\hat{8}$ that is stretched across a vast chasm—a perfect fourth—of pitch space.

But Schachter accepts Neumeyer’s readings of obvious ascending *Urlinien* in Schubertian dances (e.g., D. 779, No. 2 and D. 969, No. 7), describing these short works as akin to sections of larger works. Nonetheless, Schachter remains resistant to permitting this upward motion from becoming more than a structural oddity. For Schachter, the ascending tetrachord is sufficient to end short dances and piano miniatures, but it cannot provide the satisfactory closure that larger works require (339).

This inability for an ascending *Urlinie* to provide real and significant closure seems to quash the proposition of finding this feature in a substantive work. Schachter, for instance, seems to suggest that if the analyst identifies an ascending fundamental line within a large work, the analyst simply is not looking hard enough for the “true” descending line. I prefer a more charitable interpretation that accepts Neumeyer’s reading of a structural ascent at the end of the movement while acknowledging the lack of conclusiveness that Schachter believes it would bring to a piece. While the Neumeyer’s ascent to tonic marks the formal end of a movement, it also suggests that something is

⁵ This effect is enhanced in minor when there is a half-step relationship between $\hat{b6}$ and $\hat{5}$. Furthermore, the vast space of an augmented second between $\hat{b6}$ and $(\#)\hat{7}$ is unsuited for any structural melodic motion.

left unfinished. In this way, the music projects the feeling of desire, as experienced by a virtual agent.

In Beethoven's Piano Sonata in A major, Op. 101, i, the fundamental line ends with a climb to tonic that is reached in m. 77 (Example 3.3). This upward motion is reiterated in the final two measures as the melody ascends $E_6 - F\#_6 - G\#_6 - A_6$. Although it reaches a tonal goal, this unconventional background structure leaves things unfinished, with the rising melody akin to the rise of an interrogative vocal inflection. Although the movement draws to a formal close, the listener is left with the thematized feeling of unrealized desire.

Etwas lebhaft und mit der innigsten Empfindung:
Allegretto ma non troppo.

5 *poco ritard.* - - - *a tempo*
cresc. - *mf*

11 *dim.* *p* *cresc.* *dim.*

17 *cresc.* *p* *cresc.* *p*

23 *cresc.* *sf* *p* *espressivo e semplice*

29 *pp*

Example 3.3. Beethoven, Piano Sonata in A Major, Op. 101, i, mm. 1-34

35

41

46

51

56

61

Example 3.3. Beethoven, Piano Sonata in A Major, Op. 101, i, mm. 35-66

67

73

79

85

91

97

* *ritar* - - - - - *dan* - - - - - *do*

Example 3.3. Beethoven, Piano Sonata in A Major, Op. 101, I, mm. 67-102

Yearning

Two opposing gestural types comprise the first movement of Op. 101 (Example 3.3): an outward motion and an inward motion. These are present in the very beginning of the work, since the rising of the melody in mm. 1-2 has a descending complement in mm. 3-4. Looking closer, we see that this opposition is further present in each two-bar unit. The opening melody ascends a sixth (G₄ to E₅) before falling a fourth (E₅ to B₄). Similarly, its response begins on the anacrusis to m. 3 and ascends an octave (A₄ to A₅) before journeying downward through mm. 3-4. The left hand acts in kind, although reversing its direction. The lower voice descends stepwise in mm. 1 and leaps up into m. 2. The lowest voice then drops (A₃ to C₃) from m. 3 to m. 4 and steps upward to conclude the phrase.⁶

Hatten describes the first in the gestural pair as “yearning” (1994: 57). This gesture is indexical by virtue of its teleological nature. One does not simply yearn, but *yearns for* sometimes. Yearning represents a specific type of indexical gesture in which the agent engages in a more active role in its attempt to draw closer to a virtual object. This contrasts with that virtual observing agent who gazes upward at a firework soaring above the Venetian canals in Chopin’s Barcarolle, Op. 60 (Kopp 2014). There, the connection that draws together the virtual agent and the virtual object is suggested by the indexicality of vision. The agent’s gaze is pensive and relaxed. Yearning, however,

⁶ This contrapuntal analysis disregards E₂ in mm. 1-2, treating it as a fifth voice situated beneath the four-voice texture that does not properly participate in the counterpoint. The proper tenor voice in m. 1-2 unfolds through compound melody, ascending B₄ - E₄ - F₄ in the right hand on the second division of each beat. The location of the tenor voice is seen more clearly in mm. 3-4 when it is placed in the left hand (F₄ - E₄ - D₄ - C₄ - A₃ - G₃).

abandons this passivity and expresses a more active form of agency that seeks to draw together and unite the virtual agent and its desired virtual object.

Nowhere is this connection between agent and object more pronounced than in the yearning that is embodied as a reaching gesture. With the arms extended outward from the body, the agent *literally* draws nearer to the desired virtual observed object. This movement-toward the object of desire distinguishes it from more passive indexical gestures such as the contemplative gaze.

I describe the yearning and reaching agent as a “virtual observing agent,” because, despite the action-oriented nature of the gesture, it is strongly characterized by both its indexicality and its perspective that suggests observation. Given the implied passivity of the word “observation,” I will defend a more active type of “engaged observation.” The observer’s eye remains fixed on the goal (i.e., virtual object of desire), while the virtual observer’s body appears to actively move toward it.

Yearning as Performance Gesture

We recall Cox’s mimetic hypothesis which holds that our perception and conceptualization of observed movement is understood through a sort of off-line physical mimesis. Cox explains that “[p]art of how we understand others is by imagining performing the observed actions [of others]” (2011: [9]). According to this principle, when we witness the physical movements of another, we ourselves run a mental simulation of that action as part of our comprehension of it. Studies (Iacoboni et al. 2005, Mukamel et al. 2010) that may further support his hypothesis have compared action-

performing subjects to action-observing subjects and have found homologous neural activation within the areas of the frontal and parietal lobes that are responsible for motor action. Although this mental simulation may be below the level of consciousness, the evidence suggests that to some extent we understand an action by imagining our engagement in that action.

Cox (2011: [9]) takes this a step further and applies it to music and our perception of its motion. He continues:

Part of how we understand music is by imagining
performing one or more of the following:

- a) the observed sound-producing actions (intra-modal MMI [or mimetic motor imagery]),
- b) analogous sound-producing actions (cross-modal MMI)
- c) other analogous exertions (also cross-modal MMI)

While we have extensively explored analogous exertions (e.g., gazing) in this dissertation, we have not significantly discussed *sound-producing* actions. The latter describe actions or gestures that have a sonic component to them. Recall the opening of Beethoven's Fifth Symphony from Chapter 1 in which we heard Fate's impetuous knocking. To some, this famous motive resembles the action *and sound* that one makes while rapping one's hand forcefully and repeatedly against a hard surface. Whether or not

this was Beethoven's intention, it is not difficult to understand the connection between the music and the virtual perceived action.

Our understanding of music is similarly influenced by performance gesture. Certainly all classically trained musicians have at least some experience at the piano, at least knowing the basics of its structure and mechanics. We know that loud sounds require a forceful depression of the keys, while soft sounds require delicate, fluid motion of the wrists and fingers. The location of the keys also influences the performance gestures of the pianist. The middle register is within easy reach of the hands, while the extremes of the high and low registers require the arms to be stretched outward from the body. Thus, according to the mimetic hypothesis, listeners perceive and conceptualize the music of an instrument they are familiar with by imagining playing the piece along with the performer.

Yearning and Yielding

Returning to the Sonata in A Major, we see that the implied physical movement and its associated affect of the expanding and contracting contrapuntal wedge is felt in the performing gestures of the pianist. The outward motion of the voices requires the extension of the hands and arms away from the center of the body. The contraction of the wedge is accomplished with the inward motion of the hands and arms toward middle C. We may also hear this contrapuntal wedge as the breathing of an agent. A long inward breath fills the body, expanding the lungs and broadening the chest in mm. 1-2. It is then

expelled through the complementing phrase (mm. 3-4) as the lungs contract and the shoulders lower to a more restful position.

Hatten, too, connects yearning to the indexical act of reaching. He writes:

One can easily imagine how “yearning” might correlate with “upward” motions, since upward motions are iconic with “reaching,” and “reaching” relates to yearning through metaphors such as “reaching for a higher existence.”

(Hatten: 57)

He later expands:

[Yearning] is conveyed in this movement [of the sonata] through two kinds of processes: extended rising lines or step progressions [...] and immediate upward chromatic moves.

(100)

Fitting this affect to the opening gesture in m. 1, we see that the rising melodic line, the broadening textural wedge, the increase in dynamic intensity, and the pianist’s outward-moving hands represent the desire of a virtual agent—the entity to which we necessarily ascribe this emotion.

“Resignation” is Hatten’s opposition to the above-mentioned striving gesture. He describes the process of resignation as accomplished through “yielding,” whereby the music slows its path toward its goal, whether cadential closure, dynamic climax, or the realization of a pitch tendency (1994: 57).

“Resignation” could then attach to “descending” motions, through similar derivation (as in “lapsing from an implied ascent”), or simply by being considered oppositional to “yearning.” (57)

He identifies the cadential deferrals found in mm. 19-21, mm. 21-22, and mm. 22-23 ($V_4^6 - V_2^4 - I^6$) as examples of this affect (57-8). The desire for thematic closure and tonal resolution in the subordinate key area is created both through syntax (i.e., its position in the exposition) and the cadential expectation of the cadential 6/4. The cadence in each instance is undercut as the bass slips down from $\hat{5} - \hat{4}$ (chord-root to chord-seventh), necessitating a further resolution to $\hat{3}$ in the bass, as I^6 . With the bass failing to achieve proper closure to $\hat{1}$, the cadence is forced to yield, like an airplane that defers its landing because its approach is either too high or too low. However, after circling the runway a number of times, the bass touches down ($\hat{5} - \hat{1}$) in mm. 24-25.

The movement’s many cadential failures do not go unnoticed by Kevin Korsyn, who characterizes the piece as exemplifying a “rhetoric of evasions” (1983: 24). Korsyn, too, cites the numerous deceptive cadences throughout the movement and further notes the absence of a cadence in the development section (mm. 35-54). To this point, the through-composed nature of the exposition and the lack of articulated formal junctures connecting the major sections would make a developmental cadence disconcerting; continuity, in the sense of uninterrupted flow, is thematic for this movement. Yet amid this hinge-less formal structure, the movement abounds with deceptive cadences, places

where the listener is robbed of resolution. The quantity of these potential endpoints or musical “might -have-beens” in this movement establishes resignation as the complementary pole in the yearning-yielding thematic opposition of the work.

Yielding as Performance Gesture

The deferral of closure and flouting of musical expectations in Beethoven’s Op. 101, i, has a significant corporal element that can be felt physically by listeners. In fact, to my ears the *feeling* of failed resolution precedes my recognition of the harmonic and melodic elements that comprise the cadential failure or deferral. I can identify the progression V – vi as a deceptive cadence a split-second before consciously hearing $\hat{6}$ in the bass on the second harmony. It is not uncommon for my mouth to widen slightly at the harmonic undercutting or to feel my abdomen contract in response to the discomfort of the cadential evasion. I suspect that my reaction is not unique.

When teaching the introductory course to music theory, I similarly teach my students to recognize the distinction between V and I by experiencing the musical expectation through their own bodies.⁷ They begin by standing balanced on two feet as I play tonic (I). I instruct the students to begin transferring their weight to the right foot as I play a predominant (ii⁶ or IV). I then progress to V, at which point they are standing on one foot, wobbling and often contorting the upper body to achieve balance. Finally, with

⁷ This pedagogical approach draws inspiration from Alexandra Pierce’s *Deepening Musical Performance Through Music: The Theory and Practice of Embodied Interpretation* (2007). Pierce explores the way that the body’s interaction with gravity influences musical expression (9-36). As in my classroom exercise, she equates the tonic with physical balance, writing that “balanced alignment is the body’s tonic chord” (15).

a grand gesture of my hands, I play tonic (I) and the students shift their weight back to center as their other foot returns to the floor. This pedagogical exercise enables the metaphorical tension and release of the sonic V – I to be experienced bodily.

We return to this exercise later in the semester when I introduce the deceptive resolution (V – vi). As before, the students stand on two feet for tonic, shift their weight to the right for the predominant, and flail their arms while standing on one foot for the dominant. Once again, as I move my hands to the piano to play the next chord the students begin to shift their weight back to center. But instead of tonic, I play the submediant (vi). This elicits a number of groans and laughter at the deception; students sometimes have to use their desks to catch themselves because upon hearing the “wrong” chord, they lift the foot that they are falling toward. This helps the students to understand the “deception” created in the deceptive cadence. The predominant and dominant create an expectation that is undercut by the step-wise ascent of the bass. This links the feeling of non-satisfaction with this cadential progression in the ear, mind, and body of the student. By conceptualizing the progression through an additional modality, I find that the students come to possess a better understanding of harmonic syntax.

The performative aspect of the deceptive cadence in Op. 101 is another point for consideration. The inward motion of the hands and the contraction of the contrapuntal wedge brought on by the deceptive cadence is intensified by the inclusion of vii^{o7}/vi in m. 6. The upward slipping of the bass (E₃ – E#₃ – F#₃) and step-wise descent of the highest voice (E₅ – D₅ – C₅) counteract the expanding wedge in the previous measure. The harmony marks the onset of an expressive reversal that corresponds to the

contrapuntal reversal. This is a moment when yearning gives way to yielding.

Furthermore, the chromatic ascent of the bass in m. 6 emphasizes the motion of the left hand that is drawn inward toward the body as the right hand descends in kind. Although the local tension of $\text{vii}^{o7}/\text{vi}$ evaporates in the resolution to vi , a deeper level of tension is brought about by the harmonic undercutting by the submediant in place of the expected tonic. The contrapuntal motion of the deceptive move draws the pianist's hands and arms inward, perhaps giving the appearance of self-embrace at this poignant moment.

Beethoven adds emphasis to the affective quality of the progression by setting it with a fermata, permitting both the listener and pianist an extra moment to feel the expressed affect.

The bodily tension of the inward-moving hands is found earlier in Beethoven, in the composer's Piano Sonata in C Minor (*Pathétique*), Op. 13 (Example 3.4). The first movement begins by placing the hands and arms of the pianist in a slightly awkward position to the left of the body. Notice the fifth finger of the right hand on middle C. The hands are stretched open with the left hand spanning an octave ($\text{C}_2 - \text{C}_3$) and the right hand reaching a sixth ($\text{Eb}_3 - \text{C}_4$). The texture is thick and the register is ponderous, requiring seven fingers simultaneously to sound two tightly-stacked minor triads in the lower register. As the dark tones of this opening harmony evaporate, the texture thins and the hands expand the tonic triad outward ($\text{i}^6 - \text{passing } 6/4 - \text{i}$), as if they were giving the piece room to breathe. This outward contrapuntal wedge, however, is soon undercut with the interruption of vii^{o7}/V on beat three. Kofi Agawu identifies this interruption as a mark of the sensibility found notably in C.P.E. Bach (1991: 42). The upward augmented fourth

leap of the bass (C_3 to $F\#_3$) is a particularly jarring reversal of affect; as the bass descends downward from $\hat{3}$, it is given only a sixteenth of a beat to sound $\hat{1}$ before being yanked upward by a tritone. The fully-diminished seventh chord that it supports resolves to its prescribed resolution of V on beat four, yet a sense of tension persists.

Grave.

i $i^6 p^4$ $i vii^{o7}/V$ V

Example 3.4. Beethoven, Piano Sonata in C Minor, Op. 13, i, m. 1
(Vienna: Universal edition, ed. Heinrich Schenker, 1918-21)

This discomfort is not only heard in the prolonged instability of V, but it is also felt in the body of the pianist, whose hands have come together at the center of the body. The dominant triad is voiced as snugly as possible with $G_3 - B_3 - D_4$ fingered 1 in the left hand and 1, 3 in the right hand, respectively. In this position, the arms of the pianist are drawn uncomfortably together with the elbows perhaps touching the sides of the body and muscle tension felt in the raised shoulders. These movements together create an unsettling self-embrace that is felt by the performer and parallels the unsettling feeling evoked by the harmony ($vii^{o7}/V - V$), tight texture, minor mode, and contrapuntal reversal.

Having discussed both the connection between imagined movement and its relationship to performance movement, we are ready to conclude the analysis of the first movement of Op. 101. We recall that the yearning and yielding are manifest in the contrapuntal wedge and in the outward and inward performative gestures that create a back-and-forth between tension and relaxation. The ascending *Urlinie* of the background described by Neumeyer (1987: 300-01) is reified in the foreground as the upward melodic ascent in mm. 101-02. Hatten notes the poignancy of this upward gesture with its *galant*-style delayed resolution of the leading tone to beat two of the movement's final measure (1994: 103). This is where Schachter's observation of the inconclusiveness of the ascending *Urlinie* becomes apparent. Schachter writes:

These readings [provided by Neumeyer] contain upper-voice structures that would normally characterize sections rather than whole compositions. After all, one can all encounter pieces with "incomplete" *Urlinien* (notably $\hat{5} - \hat{4} - \hat{3}$) or pieces that end on V instead of I (also with incomplete *Urlinien*, ending on $\hat{2}$). (1996: 339)

If the ascending *Urlinie* is incomplete, it is fitting, then, that it is found in the first movement of this thematically-integrated sonata cycle. In this way, the true structural closure of the descending line that Schenker and Schachter require is deferred. We leave the agent with a sense of hope, although its ultimate desire is still unfulfilled.

The final melodic ascent of the movement (mm. 101-02) is a reiteration of the *Urlinie* that previously closed the recapitulation; therefore, it can be seen as possessing those same expressive qualities. The affect of this final rising gesture is one of yearning or reaching upward, an observation made explicit by Hatten (1994: 57). We can imagine a virtual observing agent extending its arms outward and upward with each rise of the structural line. The effort of the striving gesture is apparent in the retardation of the leading tone in the final measure. This is not the rocket gesture (Agawu 1991: 30), infused with an abundant energy that propels it rapidly upward. Rather, this ascent is earned in each ascending step as it moves toward transcendence.

The contrapuntal expansion that concludes the movement marks the point of greatest textural breadth. Here, the affect of striving experienced by the virtual observing agent is felt in the pianist's reaching performance gesture, as the hands and arms extend outward from the body. The highest and lowest voices are separated by a five-octave span (A_1 to A_6) while four octaves of negative space are situated in between (A_2 to A_5). The outward extension of the hands is a clear realization of the reaching gesture associated with the yearning expressed in the work.

Although the direction of the pianist's reach is different from that of the virtual observing agent, the effect (and the affect) are the same. With the hands separated to registral extremes, the performance gesture is also focused outward from the body. Together, we find that the expanding counterpoint of the *Ursatz* and *Urlinie* is realized in the hands of the performer and projected onto the virtual observing agent as a *yearning indexical reach*.

Part II

Observing from Multiple Perspectives

To this point, we have explored virtual objects within the diegesis that are observed by the virtual observing agent from a single perspective. For instance, in Chopin's Prelude in B Minor, Op. 28, no. 6, we recall that the virtual agent was in a distinctly observational role, rendering it a virtual observing agent per our definition. This virtual observing agent performed visual tracking gestures as though it were following the movement of a physical object rising away from it and falling into the distance. I said that although the marked melodic contour occurred numerous times throughout duration of the piece, these gestural reiterations and fragmentations were thematic transformations. In other words, the temporal pacing of the actual piece did not necessarily align with the temporal pacing of its representation. We could imagine the piece depicting the gazing gesture as though an agent were performing it with the music in the first two measures, but I suggested that the presence of the gesture later in the work was for emphasis of affect and thematic continuity. Yet despite the variations that the gesture was subject to, its singular identity remained the same. This led us to reason that the piece represented an agent gazing at a virtual object from a single perspective.

The idea of an agent taking on a single perspective within a work is appealing to an embodied theory of musical agency, because it suggests the persistence of the single subjectivity of a unified agent. Just as a photograph captures the vision of the photographer from a particular perspective, the Prelude in B Minor represents the virtual

observing agent viewing the diegesis, namely the virtual object, from a similarly fixed position. We understand that the listener experiences the music through the “eye” of the virtual observing agent in the same way that the image captured by a photograph is that seen by the photographer peering through the lens of the camera.

The analogy that likens the virtual observing agent to the photographer and camera lens can be seen by looking at Ansel Adams’s *Redwoods, Bull Creek Flat, California* (1960). This work by the famed American photographer presents an image of the towering flora in California’s Humboldt Redwoods State Park. The photograph offers the viewer a single perspective from a camera (and likely Adams’s own eyes) positioned a few meters from the tree line. The image captures the viewer, bringing her into the diegetic space constructed by the photograph. We view the image of the redwoods as though we were present in the forest where the camera snapped them. The virtual observing agent functions similarly. It brings the real audience into the virtual, diegetic space represented by the work. The listener projects herself into the diegetic space as an observing agent in the same way that the viewer of Adams’s photo embodies the photographer’s perspective and becomes present in the California forest.

One Object, Multiple Perspectives

Let us imagine that on the very same day that Adams captured this image of the redwoods, he also snapped a photo of these trees but from different angles. Suppose, further, that Adams published this in a collection under the same title. Would these photos, taken from multiple perspectives, compromise the existence of a single, unitary

agent, or would it splinter the subjectivity? Mostly likely, the single subjectivity would be preserved across the various images. Consider, in this regard, the triptych *Three Studies of Lucian Freud* (Example 3.5) by British painter Francis Bacon. In this work, the artist presents his subject from three contrasting visual perspectives. The unique geometry of the framing lines surrounding each of the three images establishes the diegetic space as consistent for the seated figure; this in turn helps orient the viewer to the virtual observing agent's perspective within the diegesis. Interestingly, each vantage point also seems to provide a distinct figurative perspective. Bacon modifies the color of the left shoe in the second frame and paints the chair with a darker shade in the third. Ernst van Alphen (1992) describes the work:

In the left-hand panel the figure is seen at some distance [depth] from the left edge of the canvas, while in the other two panels the figure is placed slightly away from the right side. In the right-hand panel the external focalizer is further away from the figure than in the other two panels. (42)

Alphen's account of the triptych interestingly acknowledges both the virtual object ("the figure") and the virtual observing agent ("the focalizer"). Within the diegetic space of the work, the virtual object remains in a generally fixed position while, as the title suggests, the virtual observing agent views it from three distinct perspectives.



Example 3.5. Francis Bacon, *Three Studies of Lucian Freud* (1969), original in color

A number of factors suggest that a single observing agent is preserved through the three panels, despite the three differing perspectives. Recall that in Chapter 1 we drew on Leonard B. Meyer's (1973: 44-79) conditions for conformance in order to explain how multiple musical ideas can join to form a single agency. These conditions of perceived conformance (Figure 1.1) are applicable here by analogy. Bacon's simultaneous presentation of the same subject, sitting in the same position, and against the same backdrop, are the most obvious contributing factors to a cohesive subjectivity. This is further supported by the consistency of color scheme, canvas size, and medium. By preserving a number of similarities among the three images, Bacon preserves both the identity of the virtual object and unifies the three visual perspectives into the single subjectivity of an observing agent.

In Chapter 1, I mentioned the slow fourth movement from Schumann's Symphony No. 3 in E-flat Major (*Rhenish*) in order to note that virtual objects can also be stationary (Example 1.5). This work represents the massive cathedral at Cologne, yet it is not representative of the structure directly. The melody, the harmony, and the timbre of the brass do not paint a musical image of the grandiose structure, the biblical mosaics in the stained glass, and the pointed steeple directly. Instead, the music replicates the physical motions that one enacts when encountering this cathedral. With this, we infer the presence of a virtual observing agent who lifts its head and eyes to take in visually the enormous structure. In listening to the music, we hear the melodic ascent as the upward motion of the head and eyes as it traces the building upward, perhaps pausing momentarily to examine the stained glass and interlocking stones on their way to the spire. *These* (deictic, indexical) observing gestures further combine with topical markers such as the solemnity of the slow tempo and the regal timbre to further characterize the indexical gesture. *This* chordal brass texture is a token of the bound style and all the religious allusions that come packaged with it. Together, *these* express the reverence that the observer experiences when encountering *this* sacred stone monolith.

Because music is evanescent and unfolds through time, it has some difficulty representing lifeless materiality. Following the suggestion of Charles O. Nussbaum (2007: 231), I hold that the musical representation of the cathedral is mediated through the movement and the senses of an observer. Nussbaum writes:

We may not be able to learn just from listening that the fourth movement of Schumann's *Rhenish* Symphony (marked "*Feierlich*," meaning "with solemn ceremony") was intended by the composer to represent the great cathedral at Cologne. But we do understand that whatever it represents is something vast, massive, and structurally intricate, with interlocking elements that demand repeated crisscrossing explorations. (231)

"How *do* we understand this?" Nussbaum (231) asks while providing his own answer:

Certain bodily movement, eye movements, head movements, and slow walks round are required to take in such a structure. (231)

Not only is Schumann's depiction of the cathedral mediated through the eyes of a virtual observing agent, the composer provides us with what I would consider an immersive experience. Nussbaum holds that the Schumann takes us on a tour of the building, portraying it to the listening "observer" from numerous points of view. Our view is drawn around the structure in much the same way as a sculpture invites us to observe from multiple perspectives. Jean Cocteau notes that "[s]culpture [...] obliges us to move round it" (1921: 12). Cone also recognizes the observer, writing that "an observing persona may be thought of as walking around a free-standing statue, so as to see all sides of it" (1974: 3). The true experience of this three-dimensional work comes in observing it from many vantage points. It is not enough to see the cathedral from the front; to more

fully appreciate it, one must travel around it, observing each detail and taking time to reflect on its majesty.

Multiple Perspectives as Expressive Doubling

The notion that one may take multiple perspectives of the same musical subject has received attention from Lawrence Kramer in his article “Expressive Doubling: Beethoven’s Two-Movement Piano Sonatas and Romantic Literature” (1988). In it, he addresses four of Beethoven’s formally peculiar piano sonatas—Op. 54 in F Major, Op. 78 in F-sharp Major, Op. 90 in E Minor, and 111 in C Minor. Kramer contends that works in the two-movement form permit “expressive doubling,” which he goes on to describe as “the configuration that results when a literary or musical work repeats the same basic pattern—whether of imagery, rhetoric, texture, or design—in contrastive versions” (176). Kramer later tightens up this description of the expressive doubling, describing it as “a form of repetition in which alternative versions of the same pattern define a cardinal difference in perspective” (1990: 22). In presenting the same basic material in two different forms, the composer provides the listener with two viewpoints of a single idea.

Kramer illustrates this in his study of the two movements of Beethoven’s Piano Sonata in F-sharp Major, Op. 78. He identifies the primary key and the dotted rhythmic figure as the single idea that binds the movements (1990: 37ff). These characteristics are seen in the incipits presented in Example 3.6. In drawing attention to these similarities, Kramer shows that the two movements are part of the same musical discourse, which

draws their unique features into direct contrast with one another. The first movement paints an idyllic scene with slow, tuneful (*adagio cantabile*) melodies and consonant harmonies. Kramer further writes that this opening movement is “expansive,” noting that both major sections have repeats (40). The second movement presents a contrast, with its aggression and instability. It opens unapologetically with immediate force (*forte*) and with the abrupt dissonance of an Italian augmented sixth chord. In comparing the two, we see the “cardinal difference in perspective” that Kramer describes (20). On the same tonal canvas (F-sharp Major), Beethoven presents two contrasting perspectives of the same dotted rhythmic idea.

The image displays two musical staves, labeled 'i' and 'ii', representing different movements from Beethoven's Piano Sonata in F-sharp Major, Op. 78. Both staves are in the key of F-sharp major (indicated by three sharps) and 4/4 time.
 Movement 'i' is titled 'Adagio cantabile.' It begins with a piano (*p*) dynamic. The right hand features a melodic line with a dotted rhythm, while the left hand provides a harmonic accompaniment with chords and single notes.
 Movement 'ii' is titled 'Allegro vivace.' It begins with a forte (*f*) dynamic. The right hand starts with a chordal texture that quickly transitions into a more active, rhythmic pattern. The left hand has a more active bass line with eighth and sixteenth notes.
 Both movements share the same dotted rhythmic motif in the right hand, which is the focus of the comparison in the text.

Example 3.6. Beethoven, Piano Sonata in F-sharp Major, Op. 78, i, mm. 1-4, and ii, mm. 1-4

If we understand the rhythmic motive common to both movements of Beethoven's Op. 78 as a musical "object," then it is not a large conceptual leap to understand its extra-musical representation as a virtual object like the cathedral at Cologne that is represented in Schumann's *Rhenish*. We hear the many ascents of the opening theme as the many curious upward gazes that an observer takes when observing the cathedral. The layering of this melody, its transfer to various instruments, and its resetting in the dominant represent distinct perspectives on the same subject. This is akin to the trope of expressive doubling in that multiple perspectives are taken of a single entity.

Multiple Perspectives in Debussy, *La cathédrale engloutie*, Préludes, Book I, no. 10

Multiple perspectives are taken by a virtual observing agent in another work about a cathedral—*La cathédrale engloutie*. For this work, Debussy draws on the myth of the Cathedral at Ys that belonged to a kingdom existing off the Brittany coast. According to legend, the principality and its cathedral lay below sea level while under the protection of a natural levee. Entry into Ys was possible only through a single locked gate whose key was held by the king. It is said that through deception, the key was stolen and turned over to an evil being—Satan, in some tellings of the story—who opened the gate during high tide. With the water rushing in at an incredible rate, Ys and its inhabitants were swallowed by the sea. Debussy's *La cathédrale engloutie* (Example 3.7) illustrates the epilogue to the story, which holds that the cathedral of the sunken Ys periodically emerges from the water to sound its bells before returning to the ocean floor.

Profondément calme (Dans une brume doucement sonore)

pp

8

4

*) Doux et fluide

8

13

pp

pp (sans nuances)

Example 3.7. Debussy, *La cathédrale engloutie*, mm. 1-15

Peu à peu sortant de la brume

16 *sempre pp* *p marqué pp*

18 *p marqué pp* *p* *marqué*

20 Augmentez progressivement (Sans presser) *

22 *più f*

26 *ff* *ff* *Sonore sans dureté*

Example 3.7. *La cathédrale engloutie*, mm. 16-30

31

35

39

43

47 Un peu moins lent (Dans une expression allant grandissant)

pp expressif et concentré

più p

pp

più pp

p

Example 3.7. Debussy, *La cathédrale engloutie*, mm. 31-51

52

56

59

63

68

pp

p

f

ff

molto

dim.

pp

Example 3.7. Debussy, *La cathédrale engloutie*, mm. 52-71

72 au Mouvement

pp Comme un echo de la phrase entendue précédemment

Flottant et sourd

75

78

81 *più pp*

Dans la sonorité du début

84 *pp*

Example 3.7. Debussy, *La cathédrale engloutie*, mm. 72-89

Debussy depicts the Cathedral of Ys from two distinct perspectives by employing the technique of expressive doubling. He presents the same musical content and its associated representation (i.e., the cathedral) in two forms. The effect is of offering the listener two distinct perspectives of the same subject. Although Kramer (1990: 20) suggests that in Beethoven's Op. 78 this "cardinal change of perspective" is figurative, the perspective of our virtual observing agent in the Debussy Prélude can be understood more literally. Because the music represents the Cathedral of Ys as seen through the eyes of a virtual observing agent, changes in perspective are not just limited to differences in conceptualization or opinion. The presence of a virtual observing agent allows a change in the virtual location of the observer within the diegetic space represented by the work. Just as the virtual observing agent walked around the cathedral at Cologne in Schumann's Symphony No. 3, Debussy is able to posit an agent who observes the Cathedral of Ys from multiple spatial locations.

The work opens with two harmonies that approach the outer edges of the keyboard; these are constructed as $G_1 - D_1 - G_2$ and $D_6 - G_6 - D_7$. The open fifths create a temporal distance as they recall the dyadic harmony of medieval parallel organum. The right-hand chord even sounds the harmonics one would hear across the empty space of the middle octaves. With this, we already receive a hint of the titular cathedral. Daniel Ericourt writes, "The first section of the piece (bars 1-15) may be taken to represent the early morning sea mists enshrouding the cathedral" (1978: 51). Debussy directs the pianist to perform with a profound sense of calmness (*Profondément calme*) at a very soft

dynamic (*pp*). With even more description, the composer writes, “*Dans une brume doucement sonore*” (In a softly sounding mist), an articulation of the haze that sits atop the waters of the Brittany coast of northern France. A pentatonic ascent of parallel fifths in the key of G begin to take shape from within the haze. Debussy sets these as inner voices. This allows the ambient cloud to persist throughout the measure, hinting at the cathedral’s emergence from out of the mist. We might hear the consecutives fifths, although soft, as stone pillars that catch our sight. Ericourt continues, describing “the rest of the piece [projects the cathedral’s] appearance (second section, bars 16-46), and eventual succumbing to the sea (third section, bars 47-89)” (1978: 51). Figure 3.3 depicts the tripartite form of the work as described by Ericourt.

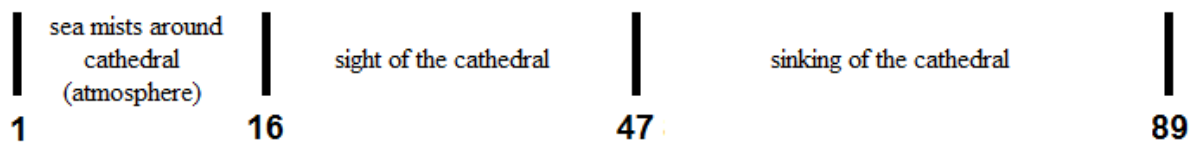


Figure 3.3. Formal layout of *La cathédrale engloutie* according to Ericourt (1978)

The cathedral begins to emerge from the fog in m. 16, where Debussy writes, “*Peu à peu sortant de la brume*” (Little by little out of the mist). We paint in our minds the scene with fog rolling through a vast landscape from which an old cathedral emerges. Monet’s *Houses of Parliament* series, which depicts the home of the British legislature on the Thames, might come to mind. The image of the cathedral is a sort of Celtic *pastorale* to which Debussy adds a touch of grandeur and mythic fantasy. One might even recall

Caspar David Friedrich's *Wanderer above the Sea of Fog*, the iconic depiction of a pensive man, standing on a rock above the sea; he casts his gaze outward onto the water and toward the mountains that peek out from the mist. As part of the landscape, the Cathedral at Ys naturally suggests an observer, and its emergence from the mist is a sight that is easily constructed in the mind of the listener. The right hand performs ascending and descending chordal structures that are imitative of the observing gestures of a virtual agent as it captures the sight of the cathedral. The observer's eyes move up and down the walls of the cathedral in mm. 16-17.

Interestingly, the observing gesture is not represented by a single melodic line, as we have seen previously. Debussy's employment of planing, as in mm. 28-40, represents a broader vision that includes the periphery of the observer's sight. Yet the melodic effect is not lost. John Clevenger holds that this triadic parallelism belongs to the melodic domain and describes planing as a thickening of the melodic line (2002: 67). To understand this, we can compare the thin stroke of a pencil to the wide stroke of a paint roller. Rather than writing with a narrow melodic line, Debussy uses planing to musically paint the image with a broad brush. In doing so, the observing agent looks past the minute details of the cathedral, and instead visually takes it in on a larger scale.

The virtual observing agent's perspective of the cathedral begins to change as it sinks back into the ocean. Debussy creates the effect of bubbling water in with a low-register major second (F#₁ – G#₁) in mm 68-70. The effervescence intensifies as the dyad oscillates in m. 71. A more fleshed-out syncopated accompaniment in the same low register begins in the following measure. The score indicates "*Flottant et sourd*" (floating

and mute), suggesting the projection of the near silence that one would hear when observing the cathedral from beneath the surface of the water. With the section marked *pianissimo* and composed with a bubbling accompaniment, the music is felt in the body just as much as it is heard in the ear. The observing gesture that viewed the cathedral in mm. 28-41 is transposed down octave to further indicate the underwater perspective of the now-engulfed cathedral.

All told, the virtual observing agent looks upon the cathedral from two perspectives that generally correspond to the formal delineations made by Ericourt (1978). Debussy uses mm. 1-15 to create an atmosphere, defining the environment of the diegesis as mythic and spacious. The middle section (mm. 16-46) portrays the cathedral emerging from the mist and coming into focus for the virtual observing agent, whose eyes trace the walls of the monumental virtual object. The cathedral descends into the sea in mm. 47-84, while the virtual agent observes from an underwater perspective. Debussy adds to this effect by indicating in m. 72 that it should sound “like an echo” (*Comme un echo*). Because the identity of the cathedral is tied to its final resting point, it is necessary to depict it musically from beneath the surface. The final measures (mm. 84-89), marked “*Dans la sonorité du début*” (In the sonority of the beginning), are a terse recollection of the beginning. It is as if the virtual observing agent were reflecting on its experience, now observing metaphorically the ever-fading memory of the lost cathedral.

Conclusion

We have moved beyond the simple gazing gesture in this chapter and have explored gestures that suggest both a yearning virtual agent and an agent viewing a virtual object from multiple perspectives. In studying yearning, we have seen that reaching gestures of the virtual observing agent can be reinforced by performance gestures of a similar nature. This corporeality has led us to further attend to the affective content associated with these indexical gestures, providing us a point of entry into the role of emotion in Chapter 4. The examination of virtual observation from multiple perspectives has allowed us to unify various indexical gestures within a single agency and has revealed the intricacy of a virtual object. This, too, leads into the next chapter, where we explore the agential entailments of perspective and narrative, as well as emotion.

CHAPTER 4

VIRTUAL OBSERVING AGENCY, EMOTION, AND NARRATIVE

Introduction

This chapter further examines the virtual observing agent and its relationship to emotion and narrative. It is a logical extension of the previous chapters because, as discussed below, there exists an intimate relationship among human expression, human movement, and human emotion. Because indexical gestures receive their individuation from the indexed objects or agents, a study of emotional transmission across this gap is warranted. The unfolding or development of emotions leads to an examination of narrative, specifically the observational and perspectival nature of a narrating agent.

I begin this chapter by drawing together research in human movement and human emotion. I then explore a typology of emotional distance that disentangles *sympathy* and *empathy*, which leads to a study of the virtual observing agent's capacity to participate as a narrator. The chapter concludes with an analysis of the second movement of Berlioz's *Harold in Italy* in which I speculate about the protagonist's emotional experience while acting as an observer to the pilgrim's procession.

From Motion to Emotion

The ability of music to arouse emotions in the listener is for many the primary reason for engaging in the listening experience. In the previous chapter, we have examined the inseparable, intermodal relationship between music and human bodily

motion. We have observed that these two aspects of the listening experience are so tightly bound together that listeners conceptualize music primarily through embodied metaphors. Lawrence Zbikowski (1997), for instance, postulates that a pitch is described as “high” because such a pitch resonates high in the head when singing while a low pitch resonates lower in the body, specifically in the chest (202-03). Similarly, a listener may describe a tempo as “quick” because quick bodily motions are required to produce the rapid rhythms that comprise it. The use of spatial and physical metaphors is not unique to music; this embodied mode of conceptualization is also used with respect to emotions. Consider the language used in the following sentences:

- 1) John has been feeling *down* since he lost his job.
- 2) The uncertainty of the situation had me feeling *tense*.
- 3) After completing her project, she felt a *weight* lifted off her shoulders.

The use of “down” to describe the emotional state in (1) is not arbitrary. Feelings of sadness are often accompanied by a lowering of both facial muscles and physical position. The frown is the most iconic image of sadness, mimicking the turned down corners of the mouth in opposition to the smile. We may also lower our bodies in times of sadness, placing our head in our hands to weep, or even descending to our knees when overcome by sorrow. Similarly, the use of “tense” in (2) captures the muscle action of anxiety. Close games are described as “nail-biters,” reflecting the tension of the jaw muscles as they engage in nervous behavior. In (3), “weight” is used to describe an undesirable burden on an individual that is relieved after a significant task has been

completed. Although often understood as metaphorical, an individual's bodily comportment may actually reflect the encumbrance caused by an actual weight. A large project that is tiring might cause us to move slowing with lowered heads and shoulders. Upon completing the project, we feel energized and stand more erect, as though a real weight has been lifted off our shoulders. In these examples, we see that our conception of emotional states is intimately tied to the bodily movements and postures that these emotions induce. Thus, the jump from human *motion* to *emotion* is not simply a creative linguistic trope; rather, it is demonstrative of an important link between body and feeling.

The first step toward uncovering the way in which a piece of music evokes a specific emotion or complex of emotions comes from understanding the nature of emotion. The analysis of this subject has a long and rich history. Plato wrote about spiritual love in *Symposium*; Aristotle developed strategies for influencing the emotions in *Rhetoric*; Aquinas disentangled passions and intellect in his *Summa Theologica*; and Descartes reduced all affective experience to a few basic emotional states. Philosophies in this vein use the human soul as a platform for emotional experience. They hold that emotions develop within the incorporeal soul, enter into our conscious experience, and manifest as a bodily response.

More modern theories of emotions, however, have viewed the *body* as fundamental to the emotional experience, not merely a consequent of it. William James (1884) famously writes,

If we fancy some strong emotion, and then try to abstract
from our consciousness of it all the feelings of its bodily

symptoms, we find we have nothing left behind, no “mind-stuff” out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains. (193)

James further emphasizes the importance of the body to the emotional process by claiming that bodily response is actually the impetus for emotion, occurring *before* the emotional experience proper. He emphasizes that

the bodily changes follow directly the PERCEPTION of the exciting fact, and that our feeling of the same changes as they occur IS the emotion. (189, italics and capitalization in original)

If we apply this theory to our experience while watching a horror film, it would have us seeing a ghastly sight, impulsively jumping and screaming, and only then feeling scared as a result of the brain interpreting our somatic reactions. For James, the emotion proper lies in the recognition of the physiological response.

An experiment by James D. Laird (1974) supports this theory and what is called the *facial feedback hypothesis*. In the study, Laird had a group of subjects manipulate the muscles in their faces to produce smiles while having them evaluate the funniness of a series of cartoons. To prevent influencing their reactions, Laird avoided emotionally charged words such as “smile” and operated under a cover—he attached electrodes to the faces of the subjects, asked them to tense and relax facial muscles one by one, and led them to believe that he was testing the electrical impulses associated with these

movements. Although misdirected by this cover story, subjects reported that they found the cartoons funnier when their own faces resembled smiles, thereby providing evidence to support James's theory that bodily response precedes and induces emotional experience.

Music and Emotion: A Physical Experience

Joseph LeDoux describes the process by which stimuli induce emotional responses in mammals in *The Emotional Brain* (1996). Writing from the perspective of a neuropsychologist, LeDoux discusses the brain and central nervous system activities involved in emotional and physiological responses. His model (1995: 223-24) recognizes the important roles of the thalamus, amygdala, and cortex. Summarized generally, the thalamus receives stimuli information directly from (most) sensory organs. It offers this information a minimal degree of evaluation before passing it to the amygdala. If warranted, the amygdala sends a signal to evoke the appropriate physiological response (e.g., recoil, jump, abdominal tension). But LeDoux holds that the sensory information received by the thalamus is also copied to the cortex, where it undergoes a slower process of greater scrutiny. Now vetted, this information is sent to the amygdala and onward to enact the appropriate physiological response.

Consider this process when listening to the second movement of Haydn's Symphony No. 94 (*Surprise Symphony*) shown in Example 4.1. The movement begins with an innocuous *staccato* melody in the strings that evokes a sense of tranquility through its simplicity and lightness. In approaching the half cadence in m. 8, Haydn

employs a secondary dominant (V/V) that emphasizes the primary dominant to which it resolves, while creating the expectation of a consequent phrase with an equally calm demeanor. This balance of form begins in m. 9 with the initiation of the same melodic and harmonic material, but in an even more subdued manner; the dynamic marking of *pianissimo* in m. 9 further reduces the intensity of the already-soft (*piano*) opening that is mm. 1-8. Having the strings play *pizzicato* and removing the second violin from the melody in mm. 9-16 also offers the melody a greater sense of lightness and buoyancy. As in the first eight-measure phrase, the melody once again proceeds to a half cadence emphasized by a secondary dominant in the penultimate measure. However, emblematic of his trademark musical humor, Haydn immediately follows the half cadence of m. 16 with the full ensemble repeating the dominant, fortissimo. Listeners unfamiliar with this piece (and many familiar with it, too) often jump reflexively at the startlingly loud sound that upsets the expectation of relaxation and breath so tastefully and delicately set up in m. 8. It is the heightened feeling of ease and tranquility of the consequent phrase that provides contrastive valence to make the fortissimo on beat two of m. 16 so jolting.¹

¹ David Huron describes the amplification of emotional response associated with surprise as “contrastive valence” (2006: 21-25, 412).

Woodwinds
Brass
Timpani

1 2 3 4 5 6 7 8

Andante *ten.* *ten.*

Violin 1-2
p

Viola
p

Cello Bass
p

9 10 11 12 13 14 15 16 *ff*

pp *ten.* *ten.* *arco* *arco* *arco*

pizz. *pizz.* *pizz.*

pp

Surprise!

Example 4.1. Haydn, Symphony No. 94 (*Surprise Symphony*), ii, mm. 1-16

The way in which Haydn's *Surprise Symphony* elicits an emotional response in the listener, however, is rather atypical. Whereas the above passage from the symphony induces an unexpected knee-jerk reaction that is below the level of consciousness, most other affective expressions in music are less charged. In fact, one might even argue that the jarring G chord in the symphony could just as easily been accomplished by a well-timed slammed door at the rear of the concert hall, and thus classify the gesture as a

frivolous and novel “effect” that is not properly musical. Most music draws us in and persuades us to experience an emotion in a more casual way than the brutish *tutti* of the symphony, which forces us to engage by instilling a genuine, even if only momentary, fear within us.

How, then, does music express emotion without the ability to smile or present a physical threat to our well-being? Aaron Ridley (1995: 75-76) proposes that musical shapes—complexes of pitch, dynamics, rhythm, articulation, timbre, harmony—resemble inflections of the human voice in a meaningful way. Drawing the two together, he describes these emotionally-charged musical contours as *melismas*. Ridley holds that in listening to music, we imitate these melismas off-line without visible action. In other words, listeners “perform” melismas in the brain, although these may not become overt physical movements. This, of course, aligns with the mimetic motor imagery described in Arnie Cox’s “mimetic hypothesis” (2001, 2011).

We see an example of this physical empathy in the opening two-note gestures of “Lacrimosa” from Mozart’s *Requiem* (Example 4.2). These measures are a well-known expression of sorrow, even when removed from the texted lamentation that follows. Focusing on the first violin, we can understand the eighth rest on each beat as a quick, sobbing breath followed immediately by wail of grief, with each dyad moving from tension to resolution. According to the mimetic hypothesis and Ridley’s understanding of melismas, the engaged listener experiences an associated emotion, perhaps sorrow, because he or she enacts the breathing and vocal gestures in the mind, engaging many of the same neural patterns as would the same physical action. Whereas the *Surprise*

Symphony induces an emotion through reaction, listeners most often feel emotion through physical and emotional empathy with the perceived emotional expression, in this case the analogous human expression of the first violins.

The image shows a musical score for two staves. The top staff is labeled 'Violin 1' and the bottom staff is labeled 'Violin 2 Viola'. Both staves are in the key of D minor (one flat) and 12/8 time. The music begins with a piano (*p*) dynamic. The Violin 1 part features a melodic line with eighth notes and slurs, while the Violin 2/ Viola part provides a harmonic accompaniment with chords and eighth notes. The score covers two measures.

Example 4.2. Mozart, *Requiem*, “Lacrimosa,” mm. 1-2

Music and Emotion: A Cognitive Experience

To this point, we have focused primarily on the way in which musical meaning and emotion are conveyed through the intermodal conceptualization of musical gestures as physical human gestures, and through the neural imitation of these gestures. But solely taking this approach overlooks a large body of research into learned musical associations. One such example is the minor mode, which is marked for tragic affect in the Classical style (Hatten 1994: 36-38). Listeners hear sadness in pieces that are in minor because the convention of the style has so firmly solidified the association between the two that it is nearly inescapable. Of course, it is possible to see the gestural motivation behind this assimilation of music and emotion; the lowered third, sixth, and seventh scale degrees are

metaphorically lower in the emotional spectrum. Other conventional musical signs have no connection at all between physical gesture and their associated emotional meaning.

In the seminal book on topic theory *Classic Music: Expression, Form and Style* (1980: 21), Leonard Ratner identifies the pedal point as an indicator of a pastoral musical setting because it is imitative—that is, it is a Peircean *icon*—of the musette or bagpipe drone associated with rustic, idyllic music. The serenity and naïveté that it evokes are not (necessarily) achieved by imagining the gestures of the droning bagpiper; instead, these expressive states are experienced through the recognition of a conventional association that binds pedal point and an idealized idyllic setting.²

The emotional reactions evoked by these types of learned associates are fundamentally different from those triggered by the unexpected *fortissimo* entrance in Haydn's *Surprise Symphony* or the sobs of Mozart's *Requiem*. For example, the C minor harmony at the beginning of Beethoven's Piano Sonata in C minor, Op. 13 (*Sonata Pathétique*), depicted in Example 3.4. The listener does not have a reaction to an affective appraisal of a stimulus that induces an autonomic bodily reaction that the mind reads as, say, tragic passion. Certainly, the *fortepiano* dynamic marking and the low register cause an unconscious affective appraisal that may cause us to react physically, perhaps with muscle tension causing a slight jump, but it does not evoke, in itself, everything that is bound up in tragic passion. Such an emotional state is the result of classical conditioning

² In Peircean terms, a pedal point in a pastoral piano sonata is an *icon* of the bagpipe drone because it sounds like (imitates) the pastoral bagpipe. The drone is a *symbol* of a pastoral setting on account of convention; the bagpipe (or a similar droning instrument) is commonly found in art and literature that depicts an idealized pastoral setting.

through stylistic convention; it comes from the stylistic semantic content within the sonority. The C minor harmony indicates a negative emotional state because of stylistic associations of the minor mode and our expectation that the first harmony of the piece will be the tonic which indicates the mode. The dense textural layering of the chord in the lower register is also a conventional sign of darkness or dread. Because these affective states are not the result of a pre-cognitive affective appraisal of a stimulus, they are not emotions in the proper sense of the word. Instead they are proclivities toward true emotions that we might think of as affective tendencies.

Psychologists have described these affective tendencies as *moods*. Unlike emotions, moods are not typically induced by pre-cognitive affective appraisal. Furthermore, they are often less acute (Lazarus 1991), more lasting (Ekman 1994), and less narrowly directed (Frijda 1994). The tragic affect of the C minor tonality of Beethoven's *Sonata Pathétique* may be *expressive of* an emotion, but in a proper sense, it can only cause us to experience a mood. When we then hear the gestural shapes within the music—the melodic contour, rhythmic patterns, and dynamic figures—we imagine our bodies performing these gestures, perhaps even enacting the associated breathing and slight muscular patterns. With the mood having lowered the threshold toward a negative emotion (e.g., tragic passion), we are more likely to actually experience this emotion as listeners, because the musical discourse further engages us and develops that emotional process.

Modes of Emotional Engagement

In describing the emotion that we hear in music, we are necessarily giving music agency. Because a non-sentient entity lacks the capacity for expression, hearing any emotion in music implies that we conceptualize the work as having the ability to reason and emote. Neural and physical mimesis permits a listener to conceptualize the expressive content of a work and co-experience that expression. More accurately, the listener anthropomorphizes the music, inferring its sentience, and then ascribes the perceived emotions to the work (or a particular passage within it.) Theorists have collated these emotions and have conceptualized them as expressed by an agency (Hatten 2004, Monahan 2013) or persona (Cone 1974, Robinson 2005) that exists *within* the work. One, then, can say that a sad passage is sad because it—the agent or persona ascribed to the music—expresses a sadness that is identifiable by the listener. Such an assertion, however, requires an asterisk to address the questions that it undoubtedly raises. What if the listener lacks the ability to understand the emotions? What if the listener inhibits the experience of the supposed “expressed” emotion? What if the listener is disinterested?

It is common for a listener to feel and embody the emotions expressed in a piece of music, yet it is not infrequent that he or she experiences emotions that are disengaged, idiosyncratic, reactive, or orthogonal to those expressed. Aaron Copland notes in his audience primer, *What to Listen for in Music* (1939 [1957]: 9-19), that a piece can exist on one of three musical planes, depending on the type of engagement chosen by the listener or afforded by the situation. One might situate Beethoven’s *Moonlight Sonata* in the background of consciousness at dinner, in the foreground of appreciation at a recital,

and under the close critical scrutiny of analysis in the classroom. Furthermore, a listener may also experience an emotion that is incongruent with the one expressed, as might be the case when one hears something objectionable. For example, the Nazi anthem, *Horst-Wessel-Lied*, is triumphant in rhythm and lyrics, yet it conjures understandably negative emotions in many Western listeners. It is necessary, then, for a theory of musical emotion to address only valid emotional experiences expressed by the music through the embodiment of a virtual agency or persona staged within the music. Although this distinction may seem artificial and inappropriately contrived, it follows the methods by which validity is circumscribed in other theories of musical interpretation that strive for intersubjectivity. Robert Hatten identifies “aesthetically warranted emotion (AWE) [as] those real, measurable emotions which are directly motivated by stylistically competent interactions with [...] those musical discourses that involve the intentionally conceived presentation, development and interaction of expressive states” (2010: 83). Such a definition removes overtly solipsistic interpretations and instead aims to endow a prototypical listener with the requisite sensibilities for meaningful engagement.

The presence of two sentient beings is necessary for there to be any type of emotional engagement in music. On one end lies the diegetic agency that expresses the emotions ascribed to the piece. This can take on a number of forms: composer, narrator, interpreting conductor, performer, virtual gesturing agent, and virtual observing agent. On the other end lies the agency that is external to the work. Although often described as a real listener, this non-diegetic agent is better understood as an abstraction of a collective subjectivity borne out of a general agreement among many listeners. This

collective subjectivity may take into account specific listeners (e.g., critics, analysts, and notable musicians), audiences contemporary with the composer, scholars, relevant social and cultural reception histories, or a mixture or subset of these.

The relationship between internal and external agencies is fundamental to characterizing the emotional experience. I describe this relationship between an internal sentient agent and an external sentient agent by delineating four *modes of engaged emotional experience*: simple cognitive recognition, empathy, sympathy, and divergent affection.

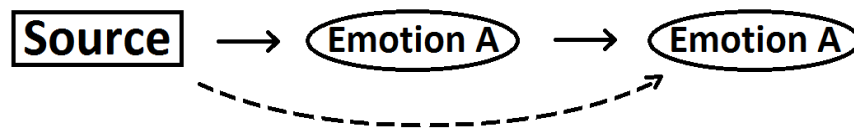
1. *Simple cognitive recognition* occurs when we as listeners cognitively recognizes an emotion that another is experiencing, yet do not become emotionally invested ourselves. This often occurs when we are indifferent to the person or agent that is experiencing the emotion first hand. When greeted with a cheery smile by coworker on a Monday morning, we might recognize the exuberance but not have the energy or desire to engage with a corresponding emotion.
2. *Empathy* means co-experiencing (feeling with) another's emotion. When we see someone smiling, we may reflect the smile on our own face, which places us in a more positive mood. Upon observing the emotional expression of another, we have a congruent emotional experience, *as if* affected by the same emotional source or stimulus (Figure 4.1a).
3. *Sympathy* is a mode of engagement that occurs when we observe the emotional expression but have a different emotional experience. Although

this expression is often in the same vein as that of the person we observe, it is *not as if* our own emotion is a direct emulation of the initial emotional source (Figure 4.1b). For example, if we were to see a close friend sobbing after having lost her job, we would be likely to feel a similarly negative emotion, yet it probably would not bring us to tears. We might comfort the friend by saying, “*I feel for you.*”³ Our own emotional experience is not nearly as palpable as hers, but it is generally in alignment with it.

4. *Divergent affection* exists when our emotional experience is markedly incongruent with that of another. For example, if a coworker were to ecstatically tell us that she received the promotion that we were also up for, we might feel dejected because we envy that person. (Hopefully, we would know to enough to dampen this expression!). This mode of emotional expression is distinct from sympathy because the emotions are particularly divergent.

³ I draw the characterization of “feeling with” (empathy) versus “feeling for” (sympathy) from Robinson and Hatten 2012: 104.

a) Empathy



b) Sympathy

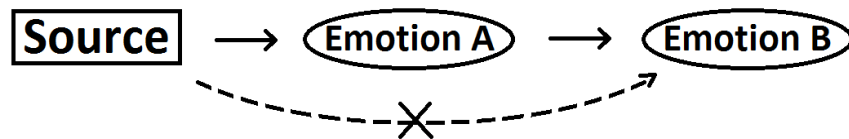


Figure 4.1. Transmission of Emotion: a) Empathy and b) Sympathy

As mentioned, there is a degree of overlap in these emotional types. Cognitive recognition, for example, is necessary for us to partake in the other three affective modes. Our reactions of empathy (feeling *with*) and sympathy (feeling *for*) involve our cognitive recognition of the source emotion that is felt in another. It is by mentally identifying another's emotional state that we are often able to feel the appropriate, corresponding empathetic or sympathetic emotion in our own bodies. Divergent emotional experiences also involve cognitive recognition, since they too are reactions to the initial emotional expression of another. For instance, we might identify the sadness felt by an antagonist while feeling a sense of happiness. It is also common for us to feel empathy and sympathy simultaneously, as often happens at a funeral. When a friend loses a parent, we express sympathy; our sadness is something that we *feel for* our friend. In this same situation, we may also experience empathy to the extent that we remember losing a

parent ourselves and *feel* grief *with* our friend. Thus, empathy involves a more direct embodiment of the emotional state of another, sympathy is more distant, and the divergent emotional experience is at an even further affective remove.

Neuroscientists Frederique de Vignemont and Tania Singer identify four necessary conditions for human empathy. They write that empathy exists if

- (i) one is in an affective state; (ii) this state is isomorphic to another person's affective state; (iii) this state is elicited by the observation or imagination of another person's affective state; (iv) one knows that the other person is the source of one's own affective state. (2006: 435)

For example, if my significant other tells me that she has recently received a promotion, and this news elicits a genuine excitement within me, it is likely that I am empathizing with her. With reference to the conditions put forth by de Vignemont and Singer, the emotion that I feel [i] is the same [ii] as the one that I have witnessed [iii] my significant other [iv] experience . By contrast, we can understand sympathy as an emotional reaction that does not meet the second condition. For example, consider the situation in which I see an undergraduate sitting in the hallway, nervously waiting to be called for her semester performance jury. Rather than beginning to shake and perspire myself, I offer a comforting smile and say, "You'll do great." Although I feel an emotion [i] at the sight [iii] of this student [iv], I *do not* co-experience [ii] her emotion. Instead, I view her expression of anxiety from an *empathic distance* that enables me to react with a different expression. The difference between the emotion observed and the one experienced is the

fundamental distinction. De Vignemont and Singer contend that the experience of empathy is modulated by four factors: the emotion itself (i.e., intensity, salience, and valence); the relationship between empathizer and target; the characteristics of the empathizer (i.e., gender, ability to understand, personal arousal tendencies); and the context of the situation (i.e., justification for emotion, interfering emotions) (2006: 437-38). All of these factors influence the empathic distance between the emotion expressed by the target and the one experienced by the observer.

The empathic distance—the degree to which the emotions of the observer and the target differ—is essential to a theory of musical emotion because emotions that lie on the surface of the music are not always those experienced by the listening agent; more exactly, these are not always expressed by the primary agency within this work. As I alluded to above, a piece of music may contain a number of virtual agents arranged into a hierarchy. Seth Monahan (2013) outlines four agent classes: 1) analyst, 2) fictional composer 3) work-persona, and 4) individuated elements (327-33). These “can appear in an array of alternate guises and/or with a number of metaphorical overlays” that Monahan calls “avatars” (348). Avatars represent more specific embodiments of agency, such the emotions of a programmatic character, the intentions of the composer, and the actions of a particular instrument part (347-53).⁴ Monahan recognizes that it is not uncommon for listeners to shift their attention from one agent class to the next (221-22). We might describe the yearning expressed by Beethoven’s Piano Sonata in A major, Op.

⁴ Although an avatar (e.g., “the pianist”) may have a real-world counterpart (e.g., Murray Perahia) the focus of any non-empirical analysis properly remains on the former.

101 (work-persona) in one sentence, while in the next sentence comments on the cunning of Beethoven—that is, a fictionalized version of him—for continually depriving the work of its expected closure. Here, we see a hierarchical arrangement of agency within a piece and our ability as listeners to shift focus from one level to the next without incident. Furthermore, because this “cunning” Beethoven is external to the interior, diegetic world of the piece, we would say that *the piece* expresses yearning rather than cunning. To measure the empathic distance between the expression of the listener and that of the piece, one would determine how the emotion of the listener compares with that of yearning.

Consider Liszt’s *La lugubre gondola I* (S. 200/1), an elegy written in Venice, perhaps as a presentiment of Wagner’s death.⁵ As seen in Example 1.4, the piece employs compound meter (6/8) to depict the buoyancy of the gondola. The accompaniment (Example 1.3) that dominates the first two sections (mm. 1-38 and 39-76) reflects the alternation between difficulty and ease in propelling the craft with an oar. We can imagine Liszt mourning at the edge of the canals and watching the gondola that holds the deceased Wagner pass with the current below him.

Treating an engaged listener as an observing agent, we can speculate about the emotional distance between this listening agent and the virtual agent within the work,

⁵ Liszt wrote in a letter to music publisher Ferdinand Taborszky dated June 8, 1885: “First of all, dear friend, will you be so kind as to go to my house with Frau von Fabry? I stupidly forgot there—in the bedroom, *not* in the *salon*—the beautiful revised copy of a composition for piano and violin or violoncello, together with the transcription of the same for pianoforte alone. The title is ‘La lugubre Gondola’ (the funeral gondola). As though it were a presentiment, I wrote this *élégie* in Venice six weeks before Wagner’s death” (1895: 473).

say, a fictionalized version of Liszt. Consider the following possibilities for the listener's emotional response:

- 1) The virtual agent's (Liszt's) expression of sorrow in the music causes the listener to feel sorrow *as if* the sight of Wagner's funeral gondola were its source.
- 2) The virtual agent's (Liszt's) expression of sorrow in the music causes the listener to feel compassion for Liszt.

The first is an example of empathy because, as seen in Figure 4.1, there does not exist a significant empathic distance between the listener and the virtual agent; the listener's emotions directly mirror those of the virtual agent. Furthermore, it satisfies the four conditions laid out by de Vignemont and Singer (2006: 435). The second scenario is an example of sympathy because the listener's emotions diverge from those of the virtual agent. While the agent's emotions incite an acceptable response from the listener, they are distinct.

Of course, the difficulty in untangling these two modes of emotional engagement is symptomatic of an inarticulate emotional language that often conflates the two. This is further complicated by the complexity of emotional expression. As mentioned above, when attending a funeral, we may feel compassion for a grieving friend (sympathy) while sharing in her sorrow (empathy). Yet we have seen that a closer examination of the emotional expression of two agents can reveal much about the empathic distance between the two.

In this section, we have placed much of our focus on the degree of engagement between the non-diegetic listening agent and an agent within the diegesis. While the former is an observing agent, it is not the *virtual* observing agent that we previously discussed. Yet the analogy between the two is helpful in conceptualizing emotional engagement. In the next section, we step into the music and examine the empathic distance between virtual observing agents and the other agents upon which they gaze.

Musical Frames and Narrative

In theorizing a virtual observing agent, one cannot help but consider its connection to narrative and, more specifically, the role of the narrator. Before continuing, it is necessary to make a distinction between “narrative” as conceived by James Litzka (1989) (and adapted to music by Byron Almén [2008]), and the “narrator” as presented in this dissertation. The former is a more broadly conceived notion of narrative that, put tersely by Seth Monahan (2013), is a “tracking of meaningful change over time” (265). However, when discussing a *narrator* here, I am referring most closely to Hatten’s “narrative agent” (2014: 226), a virtual agent who acts as a storyteller, relaying the events within the diegetic space to the non-diegetic listener. The presence of this agency distinguishes the two, representing a narrower conceptualization of narrative than the one described and implemented by Litzka and Almén.

The narrator (narrative agent) is a specific type of virtual observing agent; thus the two possess a number of common features. First, both act as lenses through which the listener experiences the virtual world. Because these agents act as conduits between the

interior and exterior spaces of piece, their accounts may exhibit various inflections, ellipses, and reordering of events. Indeed, neither conveys events necessarily with accuracy, but according to a given perspective, interpretation, and intention.

However, the narrator is distinguished from the broader class of virtual observing agents in that *the narrator always exists at the highest level of the musical discourse*. Moreover, the narrative agent is akin to storyteller speaking about characters and events from which it is conceptually detached. In literature, this virtual agent often creates the necessary narrative distance through use of the past tense. For instance, the trite opener “it was a dark and stormy night” establishes the setting as an unspecified past that is temporally removed from the narrator, the virtual agent whose perspective the reader assumes. A writer may also create narrative distance through the use of third-person pronouns such as *he, she, him, her, they, and them*, while notably avoiding the first-person pronouns *I, me, we, and us*. Again, this places the narrator at a remove from the characters and events that are found in the discourse.

Music, however, lacks a proper past tense (Abbate 1991, Chapter 2) and an equivalent for third-person pronouns to set the listener at a distance from the work. Yet music can signal the past through other means, namely quotation, repetition, and topical and stylistic allusions. For example, the recapitulation of a sonata recalls past themes in a “referential retracing of the rhetorical materials laid out in the exposition” (Hepokoski and Darcy 2006, 19). Beethoven’s String Quartet No. 15 in A Minor, Op. 132, iii similarly recalls past music, but that which is outside of the work. The slow movement

opens with imitative polyphony in the Lydian mode that draws the ear back to music of the sixteenth century.

Janice Dickensheets identifies the “bardic style” as a style topic with an implicit narrator (2012: 126-27). Unlike Beethoven’s Op. 132, iii, which references the sixteenth century from its position in the nineteenth century, works in the bardic style recall a timeless past. She writes that “[t]he effect is a musical epic retold by a poet, complete with danger, heroism, chivalry, and often, romance” (127). Unlike a simple allusion to a past style, the bard and its tale are essential to the identity of the topic, with the narrating bard often explicitly evoked through arpeggiation, itself an allusion to the primary instrument of the bard—the harp (126-27).

A narrator may also be present at the beginning of a discourse, fade into the background, and return later to the foreground. Shelley’s *Frankenstein* is a prime example. The novel begins as a series of letters written by Captain Robert Walton. Written mostly in the present tense, Walton’s narrative tells of his ongoing expedition to the North Pole and of the professor, Victor Frankenstein, that he rescues along the way. The narrating Walton writes in a letter that he greatly anticipates a story that the professor has promised to tell him, and the beginning of the central and most well-known portion of the novel is marked by a sudden change in the discourse. The perspective is transferred from Walton to Victor Frankenstein. It is here that the professor tells the tale of the monster. Although the professor speaks in the past tense, the discourse absorbs the reader

as though it were using present-tense language.⁶ Captain Walton reprises his role as narrator near the end of the final chapter when he returns to the foreground. The novel concludes from Walton's perspective as he again writes letters describing—and reflecting upon—the death of Frankenstein and his creation.

This distinction between levels of discourse is analogous to Edward T. Cone's distinction between a picture and its frame (1968: 11-31). Cone identifies two purposes of frames:

First, it separates the subject chosen for treatment from its own imagined surrounding—what I call the *internal environment*: second it protects the work from the encroachment of its *external environment*, that is, of the real time and space in which the listener perceives. (15)

Cone speaks of the frame as a boundary separating the piece of art proper from that which is outside of it. In moving forward with Cone's observations, Richard Littlefield finds this framing effect in stock literary openings, such as “once upon a time...” (1996: [3.10]). Although properly part of the literary work, the triteness of this syntactic phrase prepares the reader to enter a different discursive plane. While the narrative voice pervades the piece, perhaps speaking in the past tense and with third-person pronouns, the reader becomes engaged with the story—the lowest level of discourse—and the awareness for the narrator fades. The narrative voice reemerges when the listener is

⁶ Käte Hamburger (1973 [1957]: 64-81, as cited by Hatten, 1997) employs the term “epic preterite” to refer to the use of the past tense in foregrounded narratives. Here, past-tense verb constructions do not create their typical temporal separation; instead, the discourse is more immediate to the reader.

pulled out of the story with a similarly trite conclusion: "...and they lived happily ever after."

The Narrator in Mendelssohn, *Song without Words*, Op. 30, no. 3

Cone finds these "internal frames" in music (1968: 22), citing Mendelssohn's *Song without Words*, Op. 30, no. 3 (Example 4.3).⁷ The piece begins and ends with identical three-measure units that are detached from the body of the work. As an opening, these measures prepare by preluding, expanding the tonic triad (E major) from E₂ to B₆ through three harp-like arpeggios. The arabesque texture is like that of an ornate picture frame with sinuous, interlacing decoration that separates it from the outside world. The tightly-voiced V⁷ at the end of m. 2 supplants the opening expansive tonic and, without flourish, quickly resolves through a perfect authentic cadence that prepares the piece for its transition to the beginning proper. This cadential cliché is the not-so-subtle clearing of the throat that reigns in the textural expansion of the tonic arpeggio and prepares the entrance of the body of the piece. The desired silence follows.

⁷ Cone refers to this work as "*Song without Words No. 9 (Consolation)*" (1968: 22, emphasis mine). Although it is the third piece in the Op. 30 collection, it is Mendelssohn's ninth published in the genre, with the first six having been published in Op. 19b.

External Frame Internal Frame

Adagio non troppo

I V⁷ I

1 2 3a 3b 4

5 6 7 8 9 10

11 12 13 14 15 *p* 16

17 18 19 20 21 22

Internal Frame External Frame

23 24 25 26 27

p tranquillo

Example 4.3. Mendelssohn, *Song without Words*, Op. 30, no. 3,
framed as conceived by Cone (1968: 16-17, 22)

The body of this work begins with a visual cue, a double-bar that divides m. 3. (I use “m. 3a” and “m. 3b” to refer to each half of the measure.) Tightly-voiced harmonies, conventional 4+4 phrase structure, regular harmonic motion, and thematic continuity comprise the main part of the piece. Its chromaticism, dynamic contrast, and gazing gestures (mm. 3b-4, 7-8, 16-18) give it a feeling of hopeful determination.

The harp-like gesture of the opening returns in m. 25, and, again, it serves to separate the work from the external environment. This closing frame (mm. 25-27) is a replica of the opening frame (m. 1-3), with two notable differences. First, the resolution of the melody from the body of the work elides with the onset of the closing frame on the downbeat of m. 25. While the interface between the opening frame and the body (m. 3) is marked by silence, the body of the work bleeds into the frame in m. 25. This overlap suggests an integration between the body of the work and its internal frame. Drawing on the analogy of painting, it is as though the artist carefully chose a specific frame for the canvas and that it was an integral part of “the work.” Similarly, Mendelssohn fits the body of this work with an ornate, arabesque frame and the two become intertwined. The second difference between the internal frames exists in the final measure of each; m. 3a concludes with a beat-and-a-half of rest while m. 27 contains a fermata on the final rest. To understand this compositional choice, it is important to remember that the “outer edges” of each frame are the first and final measures of the work. Although Cone refers to the first three and final three measures of this work as “internal frames” (1968: 22), the silences that precede and follow them, respectively, are the true frames (or external frames) that delineate the so-called internal and external environments. The performer, of

course, waits for silence before playing, and it is easily earned from most audiences. The silence that follows the music, however, must often be demanded.⁸ The inclusion of the fermata on the final rest of the piece ensures that the pianist maintains a pose that signals the audience to sit in silence. The outer edge of the frame only meets the wall, so to speak, when the performer draws his or her hands away from the keyboard and gives way to the sound of the external environment.

Having both a performance frame marked by silence and an internal frame marked by the harp-like texture creates two frames for the piece, and, consequently, two narrators of different types. The silence that serves to frame the performance develops the performer as a narrator who “recalls” or “tells the story” of the work for the audience. John Rink characterizes the performer as the one who “determines the music’s essential ‘narrative’ content [...] by shaping the unfolding tale on the spur of the moment in an expressively appropriate manner” (1994: 112). Although this may be worthy of scholarly attention, this external narrator is not the one for which the theory at hand is positioned to address; therefore, I will dispense with it and pivot to the narrator evoked by the internal frames of mm. 1-3a and 25-27.

The narrative agent—topically represented as the bard—is proper to the work and is, therefore, an internal narrator. I have already noted that bardic voice evoked by the texture in Mendelssohn’s *Songs without Words* is evoked through topical allusion; the

⁸ Cone keenly observes the framing effect of silence in performance situations: “If we are members of an audience, silence should present to us a period of empty time in which *nothing* is happening. It should separate our individual and collective movements from the movement that is, for the time being, to control us all: the music. This is why the singer waits for the late-comers to be seated and for the coughing to stop. This is why the conductor poises his baton.” (1968: 16, original emphasis)

expansive arpeggios mimic the bard's harp. Although narration is essential to this topic, the narrative agency is also signaled through discontinuity, which Abbate characterizes as "disjunctions with the music surrounding it" (1991: 19). In the above *Song without Words*, the requisite musical disjunction exists in the textural differences between the arpeggios of the internal frame and the chordal texture of the work's body. Because the close voicing of the cadential progression in m. 2 and 3a reveals the texture of the body before it properly begins, a silence in m. 3a helps to effect the musical discontinuity that separates the frame from the picture. (A framing silence is not needed to create disjunction between mm. 24 and 25, because of the textural change is immediate.) This abrupt change in the musical discourse assists in securing the bardic voice, that stands apart from the lower-level discourse as a virtual observing agent or an internal narrator.

A musical disjunction creates a "shift in level of discourse," Hatten's term for a "sudden reversal [that] suggests a self-reflexive response or shift in consciousness on the part of a single agent, rather than an external agency antagonistic to the progress of a protagonist" (2004: 47). This is akin to the change in perspective that occurs at the conclusion of *Frankenstein* when there is a return to the discourse of Walton's letters. Written afterward, these letters comment on the preceding events, signaling an ascent to a higher discursive level. A similar shift occurs with the sudden reappearance of the arabesque texture in the *Song without Words*. In *Frankenstein*, the discourse begins at a higher level—one that foretells the primary discourse—and moves to a lower level—the primary discourse of the creature's genesis. An analogous shift from higher to lower occurs in the opening of the *Song without Words*.

One of the primary criticisms that Abbate levels against musical narrative is that music can *do* but not *show*. She writes:

In terms of the classical distinction, what we call narrative—novels, stories, myths, and the like—is diegetic, epic poetry and not theater. It is a tale told later, by one who escaped to the outside of the tale, for which he builds a frame to control its dangerous energy. Music’s distinction is fundamental and terrible; *it is not chiefly diegetic but mimetic*. Like any form of theater, any temporal art, it traps the listener in present experience and the beat of passing time, from which he or she cannot escape. (1991: 53, emphasis mine)

(It is important to note that her use of “diegetic” here equates to “narrative.” I discussed this meaning and the classical opposition between mimesis and diegesis in Chapter 1.) Abbate’s observation is astute—the innate immediacy of music seems inescapable, even with a supposed narrator. Mendelssohn’s *Song without Words* begins with a narrative agent, but this agent seems to disappear throughout the body of the work and only remerge in the final bars. Where does the narrator go in the interim?

The narrative agent is the lens through which the listener experiences the musical discourse that he or she comes to understand. The literary distinction between “story” and “discourse” is helpful to this metaphor. Jonathon Culler characterizes *story* as “a sequence of actions or events, conceived as independent of their manifestation in

discourse,” and contrastively defines *discourse* as “the discursive presentation or narration of events” (1981: 169-70). It can be said more plainly that the story represents actual events, while the discourse represents a recollection or interpretation of those events. Returning to the above metaphor, the events in front of the lens are the story, the lens is the narrative agent, and the discourse is the image as it is received by the viewer’s eye. Narrators can act as opaque lenses that limit the events that the viewer sees, or they can be transparent, giving the viewer a relatively unaltered view of events (although inevitably from a given perspective). In a frame story, the onset of the work’s body is often marked by an apparent dissolution of the narrator; this occurs in Mendelssohn’s *Song without Words*. Because the musical events are immediate to the listener, there seems to be no space for the narrator. It is as though the music were a story, not a discourse.

The effect of the dissolving narrator is similar to that of wearing sunglasses. When we first put on a pair of sunglasses, we are hyper-aware of the opacity of the lenses that darken the image of the world before us. Yet we know that the objects that we look at appear to us not exactly as they really are. Although the lenses alter their shading from our perspective, the objects do not actually change color. But after a short while, our awareness of the interposed lenses fades. Even though they continue to change our reception of the images before us, their darkened tint largely goes unnoticed. This narrative dissolution also occurs in film. For example, most do not recall that Miloš Forman’s *Amadeus* (1983) is actually narrated by Salieri, who appears at the beginning and ending of the film. This explains the numerous biographical inaccuracies regarding

Mozart's life. The film does not present to the viewer the story of Mozart; rather, it provides a discourse through the narrative lens of Salieri. The same can be said of the bardic narrator in the *Song without Words*. Although the narrator may disappear from view as the work shifts to a lower level of discourse, its presence is still implied and, therefore, necessary to the proper understanding and interpretation of the work.

The Narrator in Dukas, *The Sorcerer's Apprentice*

Placed at a higher level of discourse, one might assume that the narrator would not participate at the lowest level of the discourse, where the diegetic action occurs. After all, the bard in *Song without Words* bookends the work but is absent from the body. It *tells* of the main discourse but is not featured as an agent within it. However, this contrasts with the narrating Salieri, who is also a character within his historical account of Mozart's life. Similarly, in the Sherlock Holmes tales of Sir Arthur Conan Doyle, Dr. Watson is given narrative privileges while also being a participant in the narrated events.

Paul Dukas's symphonic poem *The Sorcerer's Apprentice* possesses a narrative element that is similar to the first-person perspective of Dr. Watson. Of course, its narrative component is helped along by its association with Goethe's "Der Zauberlehrling" from which Dukas derived the work; however we find that the Dukas setting is able to engender agency and perspective through its own musical means, without relying solely on the literary ballad. Having gained popularity due to its adaptation in Disney's 1940 musical classic *Fantasia* and its subsequent re-releases, the tale of an overly-ambitious apprentice is widely known.

Goethe's telling of the story begins when a sorcerer (*Hexenmeister*) leaves the side of his apprentice. The inexperienced protégée, having overheard a spell from his master, attempts to replicate an act of sorcery and animates a broom to prepare a bath. As the broom proceeds to carry the water for the bath, the apprentice begins to realize that the spell has worked too well; the anthropomorphic broom is astonishingly quick and is filling up every bucket and bowl it can find. With the tub beginning to overflow, the apprentice attempts to recall the magic words to reverse the spell, but they escape his memory. Desperate to stop the broom and defuse the situation, the apprentice takes an ax to the broom, splitting it in two. Each piece, however, regenerates itself into a new whole, with each carrying water as before. Panicked by the now-doubled efficiency, the apprentice calls out to his master to cease the enchantment. In the last stanza, we suddenly hear the words of the sorcerer—the real master—break the spell and return the broom to its original, insensate state.

Of the virtual agents (characters) that Goethe includes in his narrative (i.e., sorcerer, apprentice, two brooms), the reader is most likely to identify with the apprentice. This inclination is certainly influenced by the subject within the title of the work and the presence of first-person pronouns (i.e., *ich* and *meinem*) in the text, as well as the perspective on events that occur in the story. For example, the frenzy that is experienced as the broom begins to flood the dwelling is expressed only by the apprentice. We must remember that the actions of the broom are dutiful. Although animated, the bewitched broom is as mechanical as an automaton—unfeeling and perfunctory. It is carrying the buckets of water at a rate that might be desirable in another

situation. However, it is the apprentice—not his inanimate servant—who understands the seriousness of the situation and experiences the actions of the broom as frantic and distressing.

This perspective, which is understood through an individualized expression, is important to establishing narrative in a musical setting because it creates the necessary separation between entities within the diegesis. Language is rich in indexical (deictic) signifiers such as /you/, /she/, /this/, and /that/, words that the speaker uses to direct his or her attention toward objects or other agents. For music to point, however, it must convey this indexicality by portraying physical gestures through combinations of melodic contour, rhythm, dynamics, and articulation. Although music may take a bit more effort to clearly distinguish /this/ from /that/, it can express such a distinction by creating a separation between two virtual entities, typically two virtual agents.

In Dukas's musical setting of *The Sorcerer's Apprentice*, we hear the broom performing dutifully for the apprentice immediately following the enchantment. In mm. 67ff we hear the broom's well-known 9/8 theme set in F minor. This section is primarily diatonic and is scored lightly with woodwinds and strings to portray the nimble and effective servant broom. From the perspective of the apprentice, the animated broom is performing its task of carrying water as planned and to his satisfaction. But it is not long until the music runs amok, filling with chromaticism and an ever-thickening texture. The piece grows louder and becomes more unwieldy as the apprentice becomes frantic with his realization that he has lost control of his minion. Perhaps it is during this chaos that we first recognize the apprentice as the virtual agent with which we most closely identify,

because the music expresses his panicked emotional state; after all, the broom is mindless and acts only according to the spell. Through this individuation of the two agents and by our identifying with the apprentice's panic, the music achieves the necessary distance required to suggest a narrative interpretation.

To the same degree that Dukas's work has become part of popular culture, an essay by Carolyn Abbate (1991) entitled, "What the Sorcerer Said," has become part of the scholarly canon for musicologists and theorists. In it, Abbate draws the works of Dukas and Goethe close together as she attempts to see if music permits the same degree of narrativity in the symphonic poem that language affords the epic ballad. Abbate admits early on that she is less charitable in her identification of narrative than we might hope for music, stating that she does not hear narrative in music very often. She writes,

In my own interpretations (and my own narrative behavior)
I will interpret music as *narrating* only rarely. It is not
narrative, but it possesses moments of narration, moments
that can be identified by their bizarre and disruptive effect.
(29)

With this, she seems to be discussing the distancing effects created through a shift in the level of discourse (Hatten 1994, 2004). This enables the music to reflect or comment on the ongoing musical events in much the same way as a dramatic soliloquy in which diegetic temporality halts while a character expresses his or her inner thoughts. Abbate finds this narrative instance in the final measures of the musical work, which serve as an epilogue that stands at a remove from the events of the "story." Their separation from the

rest of the music is achieved through a pregnant pause and a distinctively slow, somewhat deflated character. For Abbate, these measures are detached from the dramatic action that comprise the symphonic poem proper and serve as the single instance of narration within the work (60).

Abbate concludes her analysis of *The Sorcerer's Apprentice* by writing that with the detachment of the epilogue comes a voice “speaking in the past tense” (1991: 60). This analogy to language follows an early discussion in the essay in which Abbate claims that music’s narrative failings are in part due to a lack of the past tense as a distancing tool (52-6). Although we discussed above that the final measures of Dukas’s work suggest a past tense, it is their novel occurrence that requires a marked shift in the level of discourse. Spoken and written languages on the other hand have conventional signifiers (i.e., verbs) to indicate tense; these verb forms are so articulate and common that they do not require disjunction to indicate tense. For Abbate, it seems that narration comes easy to language because one can write pervasively in the past tense; whereas the fundamental temporality of music and the immediacy of the performance make it difficult to listen to the music—engage with it—while remaining distant from it.

As described above, narrative implies a sense of distance while music—particularly gestural music—requires us to listen with an engaged presence. Although a seeming paradox, we must remember that music has its own means of portraying distance through indexical gestures embodied by the listener. When we point to an object across the room, we create a connection between the body and the distant object, thus drawing together the immediacy of the action and the distance of the indexed object.

When a melody ascends, we may be inclined to hear the ascent as mimetic of the upward movement of our body as we gaze or point toward an object, such a bird, traveling up into the air. Although the melodic ascent imitates the (immediate) physical movements required to track the flight, the gesture suggests that the *gesturer* is indexing and portraying distance.

Hatten (2004: 227) describes a performance technique whereby a narrative element is added through a physical gesture enacted by the performer. He proposes an instance in which a pianist relaxes her posture upon reaching the coda of a sonata. The function of a coda as an after-the-end section, in which any residual tension dissipates and thematic fragments from earlier in the work are brought to mind as though they were snapshots in a photo album, lends itself to a number of retrospective gestures. She may lift her head upward in thought or recline slightly to a tranquil position of reminiscence. These gestures are indexical in that they are made in direct reference to the music emerging from the piano at the present moment; their purpose is to draw out and emphasize an expression in the music, as though the performer were pointing her finger and saying, “Do you remember this theme?” In enacting these performance gestures during the coda, the pianist creates a disjunction between her more attentive, upright position throughout the body of the sonata and her more relaxed, reflective posture. These are analogous to the sense of involvement that we feel during the piece proper and the separation that we experience during the contemplative coda. Through an indexical gesture, the performer is able to portray to the listener the narrative element of the coda,

which simultaneously evokes the immediacy of the embodied musical experience and the distance necessary for narrative.

Regardless of the individual narrative procedure, we have seen that distance is fundamental to the expression of narrative. Abbate (1991) emphasizes the past tense as creating temporal separation in language, and Hatten (1994, 2004) speaks of a shift in the level of discourse as represented in the music and methods by which a performer portrays narrative commentary through indexical performance gestures. The representation of distance within the music that these authors discuss presents an interesting intersection with the virtual observing agent. Each features a sort of interpreter through which the musical events are relayed to the listener, through the establishment of distance by means of metaphorical pointing. By embodying this interpretive agent, the listener is given a unique perspective of this virtual world. In addition to the limited view that comes with a first-person perspective, Hatten describes narrative agency as providing a “filtered” view, in that the events within the musical discourse are seen through the eyes of a narrator who may inflect them as he or she sees fit (2004: 225-26).

This is not to say that all music involves a virtual observing agent that is narrative in nature. We must remember that narrative is an event-driven perspective that implies temporality, even if it is only through the ordering of events. Although temporal in its substance, music does not always represent a temporal narrative space that is congruent with its sonic unfolding. Moreover, this representation can be entirely void of temporality; it may be a simple vignette that describes virtual objects and inner thoughts, and it is only by necessity of the medium that these are handed to the listener as

sequenced sonic events. As Edward Pearsall keenly observes, “though [music] unfolds through time, [it] does not necessarily constitute a temporal gesture on the semantic level” (2006: 41).⁹

The Sorcerer’s Apprentice is an example in which the virtual observing agent functions as a narrator who also exists within the plot itself. Much like Dr. Watson in the Sherlock Holmes stories, the eponymous apprentice describes events as he both literally and figuratively observes them, and he exerts influence on these actions as well. Dukas’s work, in conjunction with Goethe’s text, gives clear substance to the virtual observing agent. We do not just see the events through the lens of a transcendent observer, we take on the perspective of a virtual observing agent who has the capacity to produce effects on other agents and entities within the virtual environment.

Observation and Emotion in Berlioz, *Harold in Italy*, ii (March of the Pilgrims)

Berlioz wrote *Harold in Italy* (1834) in response to a request by Paganini for a significant work for viola. The piece in four movements depicts scenes observed by the youthful Harold as he travels the Italian countryside. Berlioz admitted to drawing inspiration from the melancholy young knight in Byron’s *Childe Harold’s Pilgrimage*. In addition to the influence of Byron’s narrative poem, there is an autobiographical element

⁹ Of course, gestures by their nature require time for their unfolding; however, the temporal element of the gestural gestalt is so brief that it is insufficient to propose or contribute to a narrative by itself. David Lidov writes that a gesture is a “molar unit of motion,” implying that it is indivisible, thus inseparable from its temporal component (2005: 132).

to the work, as it was written just a few years following the young Berlioz's residence in Italy after winning the *Prix de Rome*.

Despite Paganini's love of the work, *Harold* is often measured against the yardstick of the widely-praised *Symphonie fantastique*, which premiered in 1830.¹⁰ Both symphonies are programmatic, with movements threaded with an *idée fixe*; however, they diverge in the use of this thematic device. In the earlier composition, the *idée fixe* undergoes transformations that reflect the changes in mood. But the changes in *Harold* are minimal, demonstrating the viola's distance and individuation from the surrounding music. Moreover, Harold's theme is far more simplistic (Example 4.4). Compared to the obsessive 40-bar *idée fixe* introduced in the first movement of *Symphonie fantastique*, the parallel period in eight measures denoting Harold expresses the naiveté of the young traveler freshly encountering the world.



Example 4.4. Berlioz, *Harold in Italy*, i, mm. 38-45, solo viola, first entrance of the *idée fixe*

The observational aspect of *Harold in Italy* is earned both through the work's reference to Byron's epic poem and through the contrast between the solo viola and the rest of the orchestra. Edward T. Cone notes that Harold is an observer because his theme

¹⁰ After hearing the symphony for the first time, Paganini famously wrote in a letter dated December 18, 1938: "*Beethoven spento non c'era che Berlioz che potesse farlo rivivere.*" ("Though Beethoven has died, Berlioz could revive him.") (Enrico Panzacchi 1895: 203, translation mine).

is presented in the introductory frame, outside the bounds of sonata-allegro proper of the first movement (1974: 91-93). When we identify with Harold, we establish him as a narrative agent within the diegesis who relays to us what he sees and feels. Cone further holds that the symphony is not an “objective narration”—that is, it is not simply a series of Italian scenes presented to the listener (93). Rather the “composition reports a subjective experience” through the eyes and body of Harold (93). This mediation by a narrative agent is the distinguishing element. Harold stands interposed between the audience and the procession of pilgrims, creating two levels of diegesis. Figure 4.2 shows the way in which information travels to the listener, always from the diegetic space to the non-diegetic space. When a narrator is present, the information travels within the diegetic space to a narrator who in some way modifies it (e.g., ordering, clarifying, muddying, etc.) before sending it to the listener. Because the flow of information always travels toward the narrator on its way to the audience, two levels of diegetic space are created. This is the same type of diegetic framing function that we discussed above in both the *Song without Words* and *Frankenstein*. The bottom row of Figure 4.2 depicts this framing function in the second movement of *Harold in Italy*. Here it is clear that we, the listener, do not have direct access to the “sight” of the pilgrim’s march. We learn about it through the musical narration of Harold, who describes the scene in a manner that he deems fitting for us. Like Watson and Goethe’s/Dukas’s apprentice, he is a character within the diegesis. But unlike these two, Harold’s role is more passive, as he is primarily an observer who comments upon the action and does not enact obtrusive, consequential changes within the diegesis.

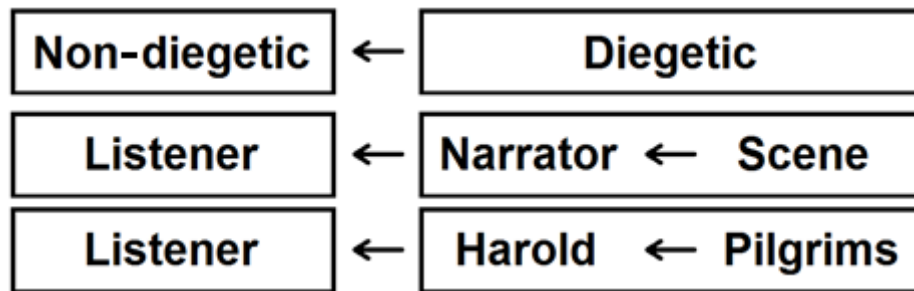


Figure 4.2. Flow of diegetic information to the listener

This narrative agency functions as a virtual observing agent in that it indexes elements of the diegesis—the pilgrims, countryside, and the church bells in the final bars. The pervasive *idée fixe* of the viola represents Harold’s presence throughout the movement. Berlioz clearly individuates this theme, demanding that the listener understand it as *standing apart* from the rest of the music. In Example 4.5, we see that although the theme appropriately harmonizes with the surrounding music, it maintains its own character. This is perhaps best heard in the solo viola’s mid-phrase entrance in m. 58. Rather than joining the rest of the ensemble at the start of a new phrase, the viola enters with the passage *in media res*, as though the virtual agent Harold unexpectedly happens upon the procession.

54 55 56 57 58 Solo. 59 60 61

Solo Viola

Violins pizz.

Viola pizz.

Cello unis.

Bass pizz. sempre

62 63 Thème de l'Adagio. 64 65 66 67 68 69

mf

Viol. Canto mf

Example 4.5. Berlioz, *Harold in Italy*, ii, mm. 54-69

It is to be made clear, however, that the *narrating* Harold is not literally represented by the viola. Instead the presence of viola represents moments when narrating agent Harold—the virtual observing agent—more directly injects himself into the musical story. If we understand the music as text, the narrating Harold is present from

beginning to end and exerts full control over the musical tale that he tells. Each melody, each timbre, each dynamic, and the organization of these musical events reflects the narrator describing the scene as he observes it. (Thus, the musical idea indicating the pilgrims might be thought of as the narrator observing, “Those pilgrims are marching along the path.”) But, of course, narration always prefixes an outwardly objective statement with recognition of the narrating agent and, in this instance, the virtual observing agent. This adds a performative aspect to the previous state, now understood as the narrating Harold stating, “*I observe* those pilgrims marching along the path.” The observing aspect of the statement is helped along by the deixis of the pronoun *those*. Continuing with the analogy, we can understand the entrance of the viola to be a more explicit first-hand account, with the narrating Harold perhaps saying, “*I have just come across* those pilgrims marching along the path.”

As both a narrator and a virtual observing agent, Harold places himself at a distance from this pilgrim’s procession that he happens upon during his travels. Although the soloist’s theme complements the surrounding harmony (suggesting favorable sympathy), Berlioz prevents it from becoming fully absorbed into the texture, thus avoiding empathetic attunement. Harold’s theme begins in m. 62 at the return of the *canto* in the second violin. While the *canto* sits atop the texture and unfolds with a relatively high degree of rhythmic activity, Harold remains detached, pacing along in half notes and whole notes. His melodic movement at the level of the measure rather than the beat is expressive of his passive and observational role. Harold’s music is thoughtful and

pensive, as if the young lad were awestruck by the world that he meets on his sojourn abroad.

Having established Harold as a narrating virtual observing agent who exists within the story that he tells, we are poised to speculate about his emotional experience. Does Harold lose himself in the experience, feeling the same emotions as the pilgrims in an expression of empathy? Or does Harold stand at an emotional remove, feeling a different complex of emotions?

For the case of empathy, we see that Harold's theme harmonizes well with the ongoing chorale, while also projecting reverence in its slow unfolding. Comparing his theme in mm. 62ff to its realization in the first movement, we notice that Harold adapts it to accommodate the duple meter of the marching pilgrims. It is as though he sings along with their tune while being sensitive to their expression of veneration.

The case for sympathy, however, seems much stronger. While he engages with the pilgrims and shares in their experience, Harold remains (somewhat) emotionally distant from them, understanding himself as a stranger in a strange land. We understand this from the manner in which Harold places his theme on top of the canto. Apart from the one-versus-many textural dynamic, the pilgrim's hymn unfolds in eight-measure phrases, divided into two-measure sub-phrases. Harold creates syncopation, both in the offbeat entrance in m. 58 and in the theme itself. Although the rhythmic augmentation of the first-movement's theme (Example 4.4) helps it to be less imposing to the chorale, it does not shed its triple meter. This creates syncopation in the hypermeter, with the

pilgrims performing in two-measure units while Harold sings along three measures at a time.

In all, it seems that Harold's expression of reverence is different from that of the pilgrims. He observes them and engages with their song, but remains at a remove. Interestingly, Harold does not use indexical gestures to create the narrative and emotional distance between himself and the virtual object that is the procession. His role as virtual observing agent is achieved instead through the lack of complete musical integration, to wit, the metric misalignment. Although the music lacks a real musical gesture of indexicality, Harold's virtual distance in the diegesis suggests that he casts an indexical gaze upon the pilgrims as he observes their sacred ritual.

Conclusion

In this chapter, we have further explored the interpretive potential of the virtual observing agent by bringing together narrative and emotion. We have found that current research suggests an intimate link between physical human movement and human affective feeling. Because musical gestures are not often indexical, evaluating the empathic distance between two agents creates the necessary space and individuation to suggest the presence of a virtual observing agent, even in the absence of indexical gestures. Shifts in the levels of discourse are also indicative of emotional distance, interpretable as musical narration. When a virtual observing agent exerts control over the ordering and interpretation of "story" events, a narrative agency emerges. By exploring this agency, we can further articulate the meaning and expression in the music.

Epilogue

I return to Chopin's Prelude in B Minor, Op. 28, no. 6 in order to briefly demonstrate an analysis and propose a line of inquiry now that we have a more complete theory of virtual observing agency in hand. In Chapters 1 and 2, I claimed that the melodic contour that we first hear in the left hand of mm. 1-2 imitates the shape of a gazing gesture. Specifically, this contour represents the motions of the head, neck, and eyes as an observer visually traces the path of motion of an object that ascends in the foreground and descends in the distance. I characterized the gesture as thoughtful and pensive, aligning it with the "tolling bells" that von Bülow heard in the right hand and with knowledge of its performance at Chopin's funeral. Without a doubt, the lethargic tempo (*lento assai*) and the minor mode contributed to this perception of gloom. I then paired these interpretations with my gestural hearing in order to paint the image of a virtual agent gazing through time toward something not readily attainable—perhaps the past or the future.

Chapter 3 more deeply explored the nature of virtual objects, specifically objects in fixed positions that were observed from multiple perspectives. Although the opening gesture of the Prelude in B Minor suggests that a stationary agent is observing a moving object, we can speculate nevertheless on its nature. Is it moving hurriedly away from the virtual observing agent? Do non-identical iterations of the contour represent different perspectives on the same object, or do they represent new objects? Furthermore, we

might ask if the virtual observing agent yearns for the object like the agent in Beethoven's Sonata in A major, op. 101, i.

Narrative and emotion were the focus of Chapter 4. Whereas the previous chapters explored the agent and the object, this final chapter looked closely at the space between them. We found that music can create a separation between an agent and an object through indexical gesture, framing, and various types of discontinuity. Although my study of Chopin's Prelude in B Minor has not revealed a strong narrative agency, I discussed the emotional expression of the virtual observing agent in detail. I did not address, however, the empathic distance between the virtual observing agent and other agents or personas that potentially exist within the work. Drawing on Monahan's (2013) agent classes, we might look closely at the emotions of the "fictional composer" to see if they differ from those of the virtual observing agent. Do other virtual agents emerge throughout the work? If so, how do their emotions differ? A theory of virtual observing agency places new analytical instruments into our interpretive toolbox. And combined with these types of questions, the theory leads us toward developing more compelling readings of a work.

The theory that I have put forth in this dissertation does not seek to radicalize listening. Rather, it provides a new method for conceptualizing music that simply furthers those inclinations that many performers, music educators, and listeners already have. The human body has long played a role in performance practice, namely in the way that performers attune their performance gestures to the shape of the music; music educators

know the worth of physical entrainment when teaching young students; and listeners who give into the impulse move to the music understand that there is physical motion implicit in every piece. Of course, gesture is more than movement; it is movement that is necessarily communicative. Gestures are natural parts of human expression that enable us to say, “Hello,” “Good bye,” and “Thank you,” without ever uttering a word. Music similarly draws on the repertory of human bodily expressions, encoding these communicative human gestures in sound and style. Properly primed, we can easily recognize these familiar human expressive elements that exist in music.

A theory of virtual agency advances the human proclivity for recognizing movement and communicative action in music. More significantly, the explicit identification of a human-like agent permits more confident speculation about the nature of gestures and their development throughout a given work. A theory of virtual *observing* agency expands upon this. Indexical gestures in music, such as pointing, gazing, and reaching, suggest the presence of a virtual agent that is individuated from other elements represented in the musical discourse. By pointing a virtual finger or casting a virtual glance, an observing agent can help define a virtual environment and furnish it with virtual objects.

The distinction between diegetic and non-diegetic spaces has featured significantly in this dissertation; respectively, they distinguish the virtual world constructed as a musical representation from the material world in which the listener resides. Musical gestures seem to lie right where these two spaces meet, having both real and representational aspects. The real listener hears and recognizes musical gestures in

the sonic material of the music, while the meaning of these heard gestures exists within the diegesis. In the latter case, gestures may take on meanings that participate in the construction of a robust virtual world.

We have seen that indexical gestures may stratify the diegetic space, creating levels of discourse and narration. Just as a real listener may judge and comment on events within the diegesis, a narrative agent—a special type of virtual observing agent—may exert influence on the way in which the “story” of the work unfolds. Virtual observing agency makes musical narration possible because indexicality (deixis) frees music of its immediacy of representation, providing the required distance between the agent and the indexed object. This permits music to talk *about* things, rather than just *perform* actions. Stated simply, (non-indexical) virtual agents can represent “this” in music, but virtual observing agents can talk about “that.” Indexicality strengthens the connection among theories of musical gesture, virtual agency, and narrativity. This dissertation has only begun to scratch the surface; a more thorough study of agency and narrative that invokes musical indexicality and indexical gesture would pay significant interpretive dividends.

Virtual agents not only perform human actions, they express human emotions. Because emotion is tightly bound together with human movement and posture, the identification of agential gesture suggests the presence of agential emotion. In Chapter 4, I noted that the indexical aspect of the virtual observing agent may modulate this emotion because it creates an empathic distance between the expressive agent with which we identify and another agent that it observes. Our virtual observing agent may augment or diminish the emotional expression that it observes in another agent. We can speculate that

emotional divergence is also possible. Parody and satire are achieved through a difference in overt expression and a more covert underlying meaning. It would be interesting to explore the intersection of these theories.

Many of the analyses that I have conducted isolate gestures within the confines of a few measures. This has allowed us to scrutinize their shapes and energetic profiles; however, it has often come at the expense of neglecting the persistence and development of the virtual observing agent throughout a work. Our examination of Beethoven's Op. 101, i, (Chapter 3) has revealed that the expressions of yearning and yielding that exist in the cadential failures of the structural foreground are also present in the background. I have demonstrated that the ascending *Urlinie* mimics the indexical reaching gesture that the agent performs while striving for a tonic goal. This has provided us with a better picture of the agent and the manner in which its gestures are manifest within the movement. From this, it appears that additional studies that track the trajectory of a virtual observing agent through entire movements could be productive.

A study of repertoire beyond the common practice period is another logical step forward for the theory of indexical gesture. While most indexical human gestures are universal and unchanging over time, their manifestation in music differs according to style. Certainly, gestures in serialism do not unfold across consonant triads, such as found in Chopin's Prelude in B Minor, yet some probably take on the same melodic shapes and rhythmic profiles to imitate human movement. We may also find a new gestural vocabulary in the twentieth century. For example, in mining Messiaen's *Vingt regards sur l'enfant-Jésus* for gestures, we do not seem to find the same physical mimesis that

occurs in Beethoven or Schumann. Perhaps these gazing gestures are buried deeper in the music, requiring a better understanding of their stylistic encoding.

With a theory of virtual agency and indexical gesture in mind, it is appropriate to point our gaze forward, toward the future of interpretation. Beyond this study lie inconceivable possibilities for engaging with music and revealing the meaning that it holds. Human empathy, however, offers us one guarantee—if we dig deep enough into music, we will undoubtedly find ourselves.

Bibliography

- Abbate, Carolyn. 1991. *Unsung Voices: Opera and Musical Narrative in the Nineteenth Century*. Princeton: Princeton University Press.
- Adams, Ansel. 1960. *Redwoods, Bull Creek Flat, California*.
- Agawu, V. Kofi. 1991. *Playing with Signs: A Semiotic Interpretation of Classic Music*. Princeton: Princeton University Press.
- Almén, Byron. 2008. *A Theory of Musical Narrative*. Bloomington: Indiana University Press.
- Austin, J. L. 1962. *How to Do Things with Words*. Cambridge, Mass.: Harvard University Press.
- Bacon, Francis. 1969. *Three Studies of Lucian Freud*. Oil on canvas, in three parts. 78 in. x 58 in., each canvas. Private collection.
- BaileyShea, Matthew L. 2012. "Musical Forces and Interpretation: Some Thoughts on a Measure in Mahler." *Music Theory Online* 18, no. 3.
<http://www.mtosmt.org/issues/mto.12.18.3/mto.12.18.3.baileyshea.php>
- Battistella, Edwin L. 1990. *Markedness: The Evaluative Superstructure of Language*. Albany: State University of New York Press.
- Buchler, Michael. 2014. "What Goes Up." Paper delivered to the University of Texas Music Theory Symposium, Austin.
- Cadwallader, Allen and David Gagné. 1998. *Analysis of Tonal Music: A Schenkerian Approach*. Oxford: Oxford University Press.
- Clater, Michelle J. 2009. "Interposed between God and Man: Agency in the Requiems of Berlioz and Fauré." Ph.D. diss., Indiana University.
- Clevenger, John R. 2002. "The Origins of Debussy's Style." Ph.D. diss., University of Rochester.
- Cocteau, Jean. 1921. *Cock and Harlequin: Notes Concerning Music*. Trans. Rollo H. Myers. London: Egoist Press.
- Cone, Edward T. 1968. *Musical Form and Musical Performance*. New York: Norton.

- _____. 1974. *The Composer's Voice*. Berkeley: University of California Press.
- Copland, Aaron. 1957 [1939]. *What to Listen for in Music*. New York: McGraw-Hill.
- Cox, Arnie. 1999. "The Metaphoric Logic of Musical Motion and Space." Ph.D. diss., University of Oregon.
- _____. 2001. "The Mimetic Hypothesis and Embodied Musical Meaning." *Musicae Scientiae* 5, no. 2: 195-209.
- _____. 2011. "Embodying Music: Principles of the Mimetic Hypothesis." *Music Theory Online* 17, no. 2. <http://www.mtosmt.org/issues/mto.11.17.2/mto.11.17.2.cox.html>
- Culler, Jonathan. 1981. *The Pursuit of Signs: Semantics, Literature, Deconstruction*. Ithaca: Cornell University Press.
- Cumming, Naomi. 2000. *The Sonic Self: Musical Subjectivity and Signification*. Bloomington: Indiana University Press.
- de Vignemont, Frederique, and Tania Singer. 2006. "The Empathic Brain: How, When and Why?" *Trends in Cognitive Science* 10, no. 10: 435-41.
- Dickensheets, Janice. 2012. "The Topical Vocabulary of the Nineteenth Century." *Journal of Musicological Research* 31: 97-137.
- Diderot, Denis. 1771. *Oeuvres de Théâtre de M. Diderot, avec un Discours sur la Poésie Dramatique*. Amsterdam: M. M. Rey.
- Eco, Umberto. 1976. *A Theory of Semiotics*. Bloomington: Indiana University Press.
- Ekman, Paul. 1994. "Moods, Emotions, and Traits." In *The Nature of Emotion: Fundamental Questions*, ed. Paul Ekman and Richard J. Davidson, 56-58. New York: Oxford University Press.
- Ericourt, Daniel. 1978. *Masterclasses in Debussy: The Piano Music*, ed. Daniel Ericourt and Robert P. Erickson, Chapel Hill, N.C.: Hinshaw Music.
- Frijda, Nico. 1994. "Emotions Require Cognitions, Even If Simple Ones." In *The Nature of Emotion: Fundamental Questions*, ed. Paul Ekman and Richard J. Davidson, 197-202. New York: Oxford University Press.
- Greenberg, Joseph. 1966. *Language Universals*. The Hague: Mouton and Company.
- Guck, Marion A. 1981. "Metaphors in Musical Discourse: The Contribution of Imagery to Analysis." Ph.D. diss., University of Michigan.

- _____. 1991. "Two Types of Metaphoric Transfer." In *Metaphor: A Musical Dimension*, ed. J. C. Kessler, 1-12. Sydney: Currency Press.
- Hamburger, Käte. 1973 [1957]. *The Logic of Literature*. Trans. Marilynn Rose. Bloomington: Indiana University Press.
- Hatten, Robert S. 1994. *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation*. Bloomington: Indiana University Press.
- _____. 1997. "Music and Tense." In *Semiotics around the World: Synthesis in Diversity. Proceedings of the Fifth Congress of the International Association for Semiotic Studies, Berkeley 1994*, ed. Irmengard Rauch and Gerald F. Carl. 627-30. Berlin: Mouton de Gruyter.
- _____. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington: Indiana University Press.
- _____. 2010. "Aesthetically Warranted Emotion and Composed Expressive Trajectories in Music." *Music Analysis* 29, nos. 1-3: 83-101.
- _____. 2012. "Musical Forces and Agential Energies: An Expansion of Steve Larson's Model." *Music Theory Online* 15, no. 2.
<http://www.mtosmt.org/issues/mto.12.18.3/mto.12.18.3.hatten.php>
- Hepokoski, James, and Warren Darcy. 2006. *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth Century Sonata*. Oxford: Oxford University Press.
- Huron, David. 1996. "The Melodic Arch in Western Folksongs." *Computing in Musicology* 10: 3-23.
- _____. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, Mass.: MIT Press.
- Iacoboni, Marco, Istvan Molnar-Szakacs, Vittorio Gallese, Giovanni Buccino, John C. Mazziotta, and Giacomo Rizzolatti. 2015. "Grasping the Intentions of Others with One's Own Mirror Neuron System." *PLOS Biology* 3, no. 3: 529-535.
- James, William. 1884. "What is an Emotion?" *Mind* 9, no. 35: 188-205.
- Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Johnson, Mark, and Steve Larson. 2003. "'Something in the Way She Moves'—Metaphors of Musical Motion." *Metaphor and Symbol* 18, no. 2: 63-84.

- Kinderman, William. 1995. *Beethoven*. Berkeley: University of California Press.
- Kopp, David. 2014. "On Performing Chopin's Barcarolle." *Music Theory Online* 20, no. 4. <http://www.mtosmt.org/issues/mto.14.20.4/mto.14.20.4.kopp.html>
- Korsyn, Kevin E. 1983. "Integration in the Works of Beethoven's Final Period." Ph.D. diss., Yale University.
- Kramer, Lawrence. 1988. "Expressive Doubling: Beethoven's Two-Movement Piano Sonatas and Romantic Literature." *Studies in Romanticism* 27, no. 2: 175-201.
- _____. 1990. *Music as Cultural Practice: 1800-1900*. Berkeley: University of California Press.
- Laird, James D. 1974. "Self-attribution of Emotion: The Effects of Expressive Behavior on the Quality of Emotional Experience." *Journal of Personality and Social Psychology* 29, no. 4: 475-86.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: Chicago University Press.
- _____. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Larson, Steve. 1997-98. "Music Forces and Melodic Patterns." *Theory and Practice* 22-23: 55-71.
- _____. 2002. "Musical Forces, Melodic Expectation, and Jazz Melody." *Music Perception* 19, no. 3: 351-85.
- _____. 2004. "Musical Forces and Melodic Expectations: Comparing Computer Models and Experimental Results." *Music Perception* 21, no. 4: 457-498.
- _____. 2012. *Musical Forces: Motion, Metaphor, and Meaning in Music*. Bloomington: Indiana University Press.
- Lazarus, Richard. S. 1991. *Emotion and Adaptation*. New York: Oxford University Press.
- LeDoux, Joseph. 1995. "Emotion: Clues from the Brain." *Annual Review of Psychology* 46: 209-35.

- _____. 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.
- Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.
- Lidov, David. 1987. "Mind and Body in Music." *Semiotica* 66, no. 1/3: 69-97.
- _____. 2005. *Is Language a Music? Writings on Musical Form and Signification*. Bloomington: Indiana University Press.
- Liszka, James Jakób. 1989. *The Semiotic of Myth: A Critical Study of the Symbol*. Bloomington: Indiana University Press.
- Liszt, Franz. 1895. *Letters of Franz Liszt: From Rome to the End*, vol. 2. Trans. Constance Bach, ed. La Mara. New York: Charles Scribner's Sons.
- Littlefield, Richard C. 1996. "The Silence of the Frames." *Music Theory Online* 2, no. 1. <http://www.mtosmt.org/issues/mto.96.2.1/mto.96.2.1.littlefield.html>
- Maus, Fred Everett. 1988. "Music as Drama." *Music Theory Spectrum* 10: 56-73.
- _____. 1991. "Music as Narrative." *Indiana Theory Review* 12: 1-34.
- Meyer, Leonard B. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
- _____. 1973. *Explaining Music: Essays and Explorations*. Berkeley: University of California Press.
- _____. 1989. *Style and Music: Theory, History, and Ideology*. Philadelphia: University of Pennsylvania Press.
- Monahan, Seth. 2013. "Action and Agency Revisited." *Journal of Music Theory* 57, no. 2: 321-71.
- Mukamel, Roy, Arne D. Ekstrom, Jonas Kaplan, Marco Iacoboni, and Itzhak Fried. "Single-Neuron Responses in Humans during Execution and Observation of Actions." *Current Biology* 20, no. 8: 750-56.
- Neumeyer, David. 1987. "The Ascending *Urlinie*." *Journal of Music Theory* 31, no. 2: 275-303.
- Nussbaum, Charles O. 2007. *The Musical Representation: Meaning, Ontology, and Emotion*. Cambridge, Mass.: MIT Press.

- Panzacchi, Enrico. 1895. *Nel Mondo della Music: Impressioni e Ricordi*. Florence: G. C. Sansoni.
- Pearsall, Edward. 2006. "Anti-Teleological Art: Articulating Meaning through Silence." In *Approaches to Meaning in Music*, ed. Byron Almén and Edward Pearsall, 41-61. Bloomington: Indiana University Press.
- Peirce, C. S. 1998 [1904]. "New Elements (*Kaiva στοιχεία*)." In *The Essential Peirce: Selected Philosophical Writings, Volume 2 (1893-1913)*, ed. the Peirce Edition Project, 300-24. Bloomington: Indiana University Press.
- Pierce, Alexandra. 2007. *Deepening Musical Performance through Movement: The Theory and Practice of Embodied Interpretation*. Bloomington: Indiana University Press.
- Plato. 1930. *Republic*. Trans. Paul Shorey. London: William Heinemann.
- Poggioli, Renato. 1969 [1959]. "The Pastoral of the Self." *Daedalus* 88 (1959): 686-89. Reprinted in *Pastoral and Romance: Modern Essays in Criticism*, ed. Eleanor Terry Lincoln, 47-60. Englewood Cliffs, N.J.: Prentice-Hall, 1969.
- Ratner, Leonard G. 1980. *Classic Music: Expression, Form, and Style*. New York: Schirmer Books.
- Ridley, Aaron. 1995. *Value, Music, and the Passions*. Ithaca: Cornell University Press.
- Rink, John. 1994. "Chopin's Ballades and the Dialectic: Analysis in Historical Perspective." *Music Analysis* 13, no. 1: 99-115.
- Robinson, Jenefer. 2005. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Oxford University Press.
- Robinson, Jenefer, and Robert S. Hatten. 2012. "Emotions in Music." *Music Theory Spectrum* 34, no. 2: 71-106.
- Saslaw, Janna. 1996. "Forces, Containers, and Paths: The Role of Body-Derived Image Schemata in the Conceptualization of Music." *Journal of Music Theory* 40, no. 2: 217-43.
- Schachter, Carl. 1996. "Schoenberg's Hat and Lewis Carroll's Trousers." In *Aflame with Music: 100 Years of Music at the University of Melbourne*, ed. Brenton Broadstock, Naomi Cumming, and Denise Erdonmez Grocke, 327-41. Melbourne: Melbourne University Press.

- Schenker, Heinrich, ed. 1918-21. *L. van Beethoven, Klaviersonaten: Nach den Autographen und Erstdrucken rekonstruiert von Heinrich Schenker*, vol. I. Vienna: Universal Edition.
- Schindler, Anton F. 1966 [1840, 1860]. *Beethoven as I Knew Him*. Trans. Constance S. Jolly, ed. Donald W. MacArdle. Chapel Hill: University of North Carolina Press. (*Biographie von Ludwig van Beethoven*. 3rd ed. Münster: Aschendorff, 1860 [1st ed. 1840].)
- Schonberg, Harold C. 1963. *The Great Pianists: From Mozart to the Present*. New York: Simon and Schuster.
- Shelley, Mary. 1818. *Frankenstein; or, The Modern Prometheus*. London: Lackington, Hughes, Harding, Mavor, & Jones.
- van Alphen, Ernst. 1992. *Francis Bacon and the Loss of Self*. London: Reaktion Books.
- Waugh, Linda. 1982. "Marked and Unmarked: A Choice between Unequals in Semiotic Structure." *Semiotica* 38: 299-318.
- Zbikowski, Lawrence M. 1997. "Conceptual Models and Cross-Domain Mapping: New Perspectives on Theories of Music and Hierarchy." *Journal of Music Theory* 41, no. 2: 193-225.
- _____. 2002. *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford: Oxford University Press.
- Zuckermandl, Victor. 1956. *Sound and Symbol: Music and the External World*. Trans. Willard R. Trask. New York: Pantheon Books.

Vita

Ian Gerg grew up in St. Marys, Pennsylvania. In 2008, he received a bachelor's degree in music education from the Pennsylvania State University, where he studied trombone with Mark Lusk and was a member of the Schreyer Honors College. Upon completing his undergraduate degree, he moved to Austin to attend graduate school, earning a master's degree in music theory from the University of Texas at Austin in 2010. He remains active as an educator, researcher, wind conductor, and trombonist.

Email: ian.wyatt.gerg@gmail.com

This dissertation was typed by the author.