Ways of listening: an ecological approach to the perception of musical meaning

[Clarke, Eric F.]

New York ; Oxford : Oxford University Press, 2005

http://books.scholarsportal.info.ezproxy.library.yorku.ca/ebooks/ebooks2/oso/2012-10-01/3/acprof-9780195151947-Clarke

Music, Motion, and Subjectivity

DOI: 10.1093/acprof:oso/9780195151947.003.0004

Abstract

This chapter discusses the ways in which sounds can specify motion, and the significance of this motion in music. Rather than considering motion in music to be metaphorical in nature, the chapter argues that sounds are capable of perceptually specifying genuine motion, but in a virtual space rather than a real space. This sense of motion is in turn one of the ways in which music brings about a strong sense of subjective identification for a listener, the sense of motion appearing to place the listener in amongst the materials of music. These ideas are illustrated and demonstrated through examples taken from pop music (Fatboy Slim), opera (a short section from Berg's *Wozzeck*), and Classical instrumental music (a passage from a Mozart String Quintet).

Keywords: **, metaphor , virtual space , gesture , Fatboy Slim , Berg , Mozart**

Having made the case for an ecological approach to music perception, and having shown how it might be applied to a specific musical example, in this and the next chapter I explore two different aspects of the way in which listeners engage with music. **Chapter 4** looks at the way in which musical materials help to determine the "subject-position," or listening perspective, that a listener adopts; while this chapter deals with a more bodily and somatic aspect of that engagement, through the sense of motion [1] that a listener may perceive in music.

Although a listener's relationship with music is primarily an auditory one, it is far from solely auditory. The interdependence between perception and action that is emphasized in ecological theory suggests that every perceptual experience will bear the trace of an action component. In the case of music, these traces are not hard to find—they are displayed overtly in the foot-tapping, head-nodding and body-swaying that are commonly observed in even the constrained circumstances of the Western art music tradition. In many other traditions, the relationship is much more dramatically obvious. There has been increasing interest in the relationship between music and the body (e.g. Lidov **1987** ; Shove and Repp **1995** ; Clarke and Davidson **1998**), and recent discussions **(p.63)** in neuroscience of so-called mirror neurons (nerve cells in the motor area of the brain that show imitative activity when a perceiver observes another individual performing specific actions) suggest the possibility of a very deep-seated perception-action coupling. Ramachandran (**2000**) has suggested that mirror neurons may be the "single most important 'unreported' … story of the decade," and that their generalized mimetic function may hold the key to understanding the evolutionary development of distinctively human culture—including language, art, and music. For at least three reasons, therefore, there is a close relationship between sound (musical and everyday) and motion: [2] (1) sound is a direct consequence and indicator of motion, as in the clatter of stones down a hillside, the distinctive pitch and dynamic profile of a passing motorbike, or the creak of a branch swaying in the wind; (2) the relationship between auditory perception and human action is encapsulated in the functions of mirror neurons or the more general sensorimotor contingencies discussed by Noë and others (e.g., Hurley and Noë **2003** ; Noë **2003**); (3) the production of musical sound requires and inevitably involves movement, from the relatively discreet movements of the vocal tract in singing, to the more manifest movements of hand-clapping and instrumental performance.

There is a long history of writing on the relationship between music and motion, dating back to classical Greek writings on music (see Barker **1989**). My approach in this chapter is to argue that the sense of motion when listening to music is an inevitable consequence of the event-detecting nature of the human auditory system; that there are some interesting questions about *what* listeners perceive as being in motion; and that the varieties of motion specified in musical sound together constitute a crucial component of listeners' perceptions of meaning in music. In short, the relationship between music and motion is a fundamental aspect of music's impact and meaning; is significantly, but not only, concerned with the perception of self-motion; and should be regarded as a truly perceptual relationship—even though the perceived motion may **(p.64)** be illusory (in the sense of being attributed to virtual, rather than real, objects).

As an emblematic example, consider the first and second subjects of the first movement of Beethoven's Fifth Symphony (Op. 67). Virtually every program note, CD liner, or record sleeve comments on the contrast between the dramatic and explosive character of the first subject material and the lyrical and flowing character of the second. But what or who is being dramatic and explosive on the one hand and lyrical and flowing on the other? Do listeners subjectively identify with the music and hear *themselves* acting out explosive or flowing actions, or do they hear these sounds as specifying the actions of *other* agents or objects? If the latter, then are these actions the abstract and metaphorical movements of musical material, or the real movements made by the performing musicians, or the imagined movements of fictional characters?

Music and Motion

The sense of motion in music has been closely associated with rhythm and tempo (most of the conventional terminology for tempo is motion-based). At the start of his historical overview of rhythmic theory, Yeston (**1976** : 1) observes that "In the broadest sense, the theory of musical rhythm has always been concerned with the elucidation of musical motion—motion that is differentiated by the durational value, pitch, or intensity of sounds, but which, at the same time, presumably exhibits certain regularities."

But the term "musical motion" is a slippery one, since it skirts around the question of what is moving, and in what kind of "space." Yeston's own perspective lies firmly within a tradition of structuralist music theory—score-based and abstract. By contrast, Shove and Repp (**1995**), who provide an important survey of motion literature relating to musical performance, offer a very different perspective on music and motion, one that focuses on the actual movements **(p.65)** of performing musicians as an important and curiously overlooked source of the sense of motion that listeners experience.

Traditionally, to explain the source of musical motion, theorists, philosophers and psychologists alike have turned to musical structure, which by most accounts is abstract. This has led some to believe that the motion heard is virtual, illusory or abstract…. Hidden from this view is perhaps the most obvious source of musical movement: the human performer. Why have so many theorists failed to acknowledge that musical movement is, among other things, *human* movement? (Shove and Repp **1995** : 58; emphasis in original)

Shove and Repp examine the ways in which the real (or imagined) motion of the performer may be specified in sound, and the ways in which the movements of both performers and listeners might be used to increase their aesthetic engagement with music. In doing so, they demonstrate that a number of German writers in the early part of the twentieth century (Sievers, Becking, Truslit; see Shove and Repp **1995**) were particularly interested in the relationship between body movement, gesture, and performance. Each of these authors developed his own lexicon of movement types, the function of which was both analytical and practical: each lexicon was intended both to reveal, and make sense of, the inner dynamics of music, and in this way to help performers to find a fluent and expressive approach in performance. Becking, whose ideas were subsequently taken up and developed by Clynes (e.g. Clynes **1983**), distinguished a number of types or styles of movement curve, and attributed these different styles of movement to the music of different groups of composers. Truslit (**1938** ; see Repp **1993**) had no interest in the composer specificity of musical motion, but was more concerned with the relationship between the acoustic surface of a musical performance and the underlying motion dynamics of the piece. He was interested therefore in the specifics of individual pieces and in discovering particular movement patterns that would help listeners and performers to understand and project the music **(p.66)** in the most effective manner. One feature of his theory is the proposal that performers' larger (more global) and smaller (more local) movements (and the mirroring physical responses of listeners, such as foot-tapping, head-nodding, swaying, and so on) reflect the organization of the human motor system into two divisions—one controlling whole body movement, and the other more peripheral limb movements. The former (the ventromedial system) is closely associated with the vestibular apparatus in the inner ear, which is responsible for our sense of balance and motion, thus suggesting a possible direct physiological link between the perception of sound and the sense of bodily motion.

Although Shove and Repp briefly consider the wider specification of motion in music (i.e., motion that is not necessarily attributable to the performer, but to some other agent, such as the listener's own motion, or the motion of some virtual agent) this is not the main focus of their review. Todd, however, has paid rather more attention to listening and the sense of motion that it can induce. In a paper concerned with the relationship between tempo and dynamics in performance, he concludes with the proposal that expression in performance has its origins in simple motor actions, that the performance and perception of tempo and dynamics is based on an internal sense of motion, and that “expressive sounds can induce a percept of self-motion in the listener”(Todd **1992** : 3549). The basis for this, he speculates, may lie in the physiology of the vestibular apparatus in the inner ear, and in the possibility that sound directly activates that part of it (the saccule) that is responsible for a person's sense of self-motion. In a subsequent paper, he notes the need to be cautious about the relationship between musical expression and physical motion, distinguishing between “purely metaphorical notions of musical motions and any more psychologically concrete phenomena that correspond to the metaphorical” (Todd **1995** : 1946). More recently (Todd **1999**), he has argued strongly for a vestibular component within the sense of musical motion, complemented by a sensorimotor element. Rather like Truslit, he has suggested that these two systems possibly **(p.67)** map onto the distinction between smaller-scale gestural motion (movements of the limbs) and larger-scale whole body movement, respectively.

Empirical evidence for the perception of motion in rhythm is provided by Gabrielsson (**1973**), who asked listeners to rate a collection of simple rhythmic figures using a large number of descriptors. The descriptors attributed to the rhythms clustered together according to three main principles: those concerned with structure, which related to the complexity of the rhythmic patterns and their meter (triple or duple); those relating to the emotional quality of the rhythms; and, finally, descriptors relating to the motion character of the rhythms, involving words such as “running,” “limping,” “flowing,” “crawling,” and so on. Given music's long history of association with a variety of physical activities (dancing, singing while working, marching, performing ceremonial actions), this finding might not seem too surprising, but the spontaneous emergence of this factor in Gabrielsson's research comes from a laboratory study in which very simple rhythms were played on a single drum. The implication is that this motion character is a pervasive and deep-seated component of listeners' responses to even quite simplified musical materials.

The Metaphor of Motion

In their different ways, Shove and Repp, Todd, and Gabrielsson all treat motion in music as a more or less concrete perceptual phenomenon. But it has been widely claimed that the sense of motion in music—like the sense of space in music—is essentially metaphorical. [3] There is, after all, no real space that musical materials inhabit (so the argument goes), and musical elements (pitches, rhythms, textures, etc.) have no concrete material existence. Therefore, since motion is a property of objects in space, the whole idea of musical motion —if taken literally—is a nonstarter. In *The Aesthetics of Music*, Scruton (1997) gives serious attention to **(p. 68)** a consideration of music's motion character from precisely the perspective of its fundamentally metaphorical character. The starting point for Scruton's argument is the distinction between sound and tone, which he identifies in three specific attributes: the distinction between what he calls the "acoustical experience of sounds" and the "musical experience of tones"; the distinction between the "real causality" of sounds and the "virtual causality" that relates tones to one another in music; and finally the distinction between the sequence of sounds and the movement of tones that listeners hear (Scruton 1997 : 19). As these distinctions make clear, Scruton argues strenuously for a fundamental distinction between the sounds of the "everyday" world and the tones of music. This crucially places listeners' motion experiences in a realm that is quite separate from their auditory experiences of the motion character of everyday objects in the "real" world—the sound of footsteps approaching, of cars passing, of balls bouncing, of bottles breaking, of water gushing, and so on. Scruton nonetheless places the sense of motion at the center of musical experience: "Whenever we hear music, we hear movement," he writes (55), and " … we *must* hear the movement in music, if we are to hear it as music" (52; emphasis in original).

His explanation for this sense of movement in music, relating as it does to an acousmatic space divorced from the real spaces of the world, is that it depends on a deep-seated metaphor:

[The] idea of musical movement is an irreducible metaphor, which can be explained only through our response to music. It is associated with other metaphors—and in particular with the metaphor of life. In hearing the movement in music we are hearing life—life conscious of itself…. (353)

Elsewhere he suggests that we "think of music as spread out in acousmatic space, where a new kind of individual is born and lives out its life: an individual whose character is constantly changing in response to the musical surroundings" (72). Scruton's approach is a consequence of the fundamental claim that musical events are **(p.69)** "secondary qualities"—not tied to, or a part of, the physical circumstances of the real world but separated from them, and capable of behaving in ways that are not constrained by the real world. His notion of movement and gesture is abstract and idealized because the entities that move (notes in a melody, or a chord within a sequence) have no reality as physical objects. Musical motion and space are metaphorical because the properties of *real* space and motion have been transferred across to another domain where they have no literal application. Despite referring approvingly to the work of Lakoff and Johnson, who argue that metaphor occupies a central position in human cognition (Lakoff and Johnson 1980 ; 1999), Scruton asserts that "all metaphors are false"(1997 : 91). By drawing a clear line between the tangible and practical world of sound, and the abstract and incorporeal domain of tone, Scruton places music firmly in a metaphysical realm and aligns himself definitively with the aesthetic tradition of music's autonomy.

But why insist upon the separation between sound and tone? The sounds of music can and obviously do specify objects and events in the world (instruments and the people who play them), and kinds of action (bowing, blowing, plucking, striking), even when the precise nature of those actions is unclear or uncertain (a person may hear striking without knowing exactly what is being struck). In this most obvious sense the sounds of music *are* the sounds of the everyday world. A listening strategy that was concerned solely with the objects and actions of instrumental performance might be rather limited (although the importance of music's instrumental character has until quite recently been seriously undervalued by both musicology and the psychology of music), but a range of other possibilities is opened up by considering the way in which musical sounds may specify the objects and events of a *virtual* environment. In visual art, an important part of the perceptual experience of looking at a painting (representational or abstract), is to see the forms and colors of the painted surface as specifying virtual objects in a virtual space with properties that are reminiscent of everyday objects and spaces—or (as in of some of M. C. Escher's **(p.70)** drawings, for example) scenes that fascinate their viewers by defying or playing with the normal rules of space. The same principle applies to hearing: sounds can specify a virtual domain that both abides by, and stretches or defies, the normal laws of physics. [4]

Gjerdingen (1999) gives an account of what he calls apparent motion in music that takes as its starting point the analogy between apparent motion in vision and in music. A well-known illusion in vision, known as the "phi effect" (Kolers 1972), can be created by placing two lights a small distance apart, switching on and off in alternation with one another. When spaced at an appropriate distance, and flashing within the appropriate range of rates, people report a compelling sense of movement from one light to the other. [5] Similarly, two notes spaced an appropriate interval apart, and alternating at an appropriate rate, will give rise to an "apparent movement" between the notes. However, Gjerdingen asserts that

if musicians—and listeners in general—concur in sensing motion in music, there is little agreement about the *nature* of the motion…. In viewing a classic visual demonstration of apparent motion, one can easily imagine an intermittently flashing light source that moves in normal three-dimensional space. In hearing the subject of a Bach fugue, however, it is not at all clear what is moving or where that motion takes place. (Gjerdingen 1999 : 142; emphasis added)

His own approach presents apparent motion as motion in a unidimensional pitch space (simply lower to higher) over time. Implemented as a neural network, his model provides a graphic output that seems to capture many of the apparent motion effects that listeners hear in the context of phenomena such as the pitch streaming of interleaved parts, scalar and arpeggiated pitch patterns, trills, and the separation or integration of parallel parts. Persuasive though the demonstrations are, they leave a basic question unaddressed—the issue which Gjerdingen himself identifies in the extract above: *what* is moving and in what kind of space does this motion occur?

(p.71) The Perceptual Reality of Motion in Music

Recall the unlabeled CD discussed at the beginning of the introduction, with its recording of someone eating crisps from a packet. The sounds of that recording specify, among other things, the crackly movements of the crisp packet, the mouth movements of the person eating, and so on. When I hear these events over a pair of loudspeakers or headphones, I am under no illusion that the packet of crisps or the person eating them are concealed somewhere within the loudspeakers or the headphones, just as when I watch a newsreader on television I do not mistakenly believe that she is actually in the set. But neither do I see the newsreader's movements, or hear the crackling and crunching of the crisps, as metaphorical events. I perceive them as perfectly real, but I also perceive that these events are broadcast or recorded, and that the actual events are not happening here and now. In the same way, the study by Warren and Verbrugge (1984) of the perception of bouncing and breaking, referred to in chapter 1 , showed that listeners could reliably distinguish between the two kinds of event on the basis of sound recordings of auditory sequences, or artificial simulations of those sequences that preserved the principal temporal properties of the originals. These recordings or simulations are not heard as specifying apparent, or metaphorical, events: they specify perceptually real events that happen not to be present.

A solution to this problem of the perceptual reality, but physical absence, of recorded events is provided by the idea of a "virtual source"—an idea developed by McAdams (1984); and it is with this idea that ecological theory provides an explanation of the perception of motion in music. McAdams coined the term "virtual source" by analogy with the term virtual image (or virtual object) in optics, where it refers to the objects and images seen in mirrors and pictures, and which occupy the virtual space behind the plane of the picture or mirror. In a similar manner, musical sounds may be organized in such a way that they specify a source that has **(p.72)** no real, physical existence. Various tricks of orchestration are an obvious example, where the impression of a "virtual instrument" that has no empirical presence can be created through the fusion of sounds coming from various actual sources. Bregman (**1990** : 460), in an extensive discussion of auditory perception organized around the central idea of "auditory scene analysis," points out that "the virtual source in music plays the same perceptual role as our perception of a real source does in natural environments." The psychological processes involved in perceiving real and virtual sources are identical, just as the experiences themselves may be equally vivid and apparently veridical—but while one refers to real objects and events in the world, the other does not. "Experiences of real sources and of virtual sources … are different not in terms of their psychological properties, but in the reality of the things that they refer to in the real world. Real sources tell a true story; virtual sources are fictional" (Bregman **1990** : 460).

In the case of a mirror, the virtual objects that can be seen in it have a lawful relationship with the real objects of which they are a reflection, move in a fashion that corresponds directly to the movements of their real counterparts (of which they are a reflection), and are described by an optics identical to that of the real world (though reversed). In a picture, film, or animation, the objects have qualities that may mimic those of the real world, and can do so very convincingly in the case of *trompe l'oeil* painting, or computer animation, but these qualities are achieved using quite different means. In a discussion of painting and drawing (Gibson **1966** : 240), Gibson pointed out that Hans Holbein, in his portrait of Sir Thomas More, achieved the sense of folding and texture on More's velvet sleeve by using pigments of different hues, while the same visual effect on a real piece of velvet is produced by differential reflectance and shadow. Gibson noted the intriguing duality of a picture—the fact that the marked surface is both an object and also a medium by means of which a viewer may perceive other virtual objects:**(p.73)**

A picture, photographic or chirographic, is always a treated surface, and it is always seen in a context of other nonpictorial surfaces. Along with the invariants for the depicted layout of surfaces, there are invariants for the surface as such. It is a plaster wall, or a sheet of canvas, a panel, a screen, or a piece of paper…. The information displayed is dual. The picture is both a scene and a surface, and the scene is paradoxically *behind* the surface. (Gibson **1979** : 281; emphasis in original)

The comparison with painting is instructive because it suggests a way of understanding both what is perceived as moving in music, and how the effect is produced. Just as spatial patterns of pigment in a painting can create a perceptual effect analogous to that produced by reflection, texture, and shadow in the real world, so music may create perceptual effects with temporal patterns of discrete pitches that reproduce, or approximate, those that we experience with the continuous acoustical transformations that are characteristic of real-world events.

Motion, Gesture, Meaning

While it seems obvious that visual information can specify motion, there is more resistance to the idea that the same might be true of sound. Certain kinds of acoustical invariants are readily accepted as specifying motion (e.g. a continuous change in left ear/right ear intensity balance or phase relation; or the pitch shift of the Doppler effect) since these directly specify real movement in real space. The experience of vivid sound tracks or demonstration sequences in surround-sound music systems or Dolby Digital cinemas are an indicator of how powerful this effect can be even when artificially generated. But a much more general possibility is that changing patterns of attack point, timbre, dynamic, and pitch have the capacity to specify motion in a virtual space—in the **(p.74)** same way that the continuous spatial displacements of visual edges, points of light, swirls and textures in computer animation do.

Transformations in loudness, timbre, and other acoustic properties may allow the listener to conclude that the maker of a sound is drawing nearer, becoming weaker or more aggressive, or changing in other ways. (Bregman 1990 : 469)

Just as the success of animation depends on the propensity of the visual system to detect movement in even quite poor approximations to the visual arrays that specify real movement, so too the perception of motion and gesture in music relies on the detection of motional and gestural invariants in sound sequences which may, in objective terms, be quite poor approximations to their real-world counterparts. An orchestral performance of the opening motif of Beethoven's Fifth Symphony may have relatively little literal acoustical similarity to the sound of a hand knocking at a door, but it is sufficient for generations of listeners (guided by suggestive program notes and CD liners) to hear the "knocking of fate."

The basic principle can be stated simply enough: since sounds in the everyday world specify (among other things) the motional characteristics of their sources, it is inevitable that musical sounds will also specify movements and gestures—both the real movements and gestures involved in actually producing music (as discussed by Shove and Repp 1995 , for example) and also the fictional movements and gestures of the virtual environment which they conjure up. As Windsor (1995 ; 2000) has discussed, acousmatic music [6] provides a context in which this is particularly striking, since the absence of any visual anchor allows free play within the virtual spaces that the music specifies. In certain respects, this is not a new idea at all: Langer (1942) wrote of the connection between musical meaning and music's capacity to convey movement and gesture, but the difference here is my claim that this relationship is truly perceptual rather than metaphorical, symbolic, or analogical. For the obvious adaptive reasons of getting around and surviving in an unpredictable environment, the auditory system is highly **(p.75)** attuned to the motion-specifying properties of sounds, and since the variety of ways in which animate and inanimate objects may move is unlimited, every musical sound has the capacity to specify *some* kind of motion (or its opposite—stasis [7]).

Sounds specify motion by means of change [8] (discrete and continuous changes of pitch, timbre, and loudness over time); the crucial question is *what* a listener hears as being in motion. Todd has proposed that the perception of motion in music is a sense of *self*-motion, as a necessary consequence of the vestibular explanation that he proposes. If musical sound directly stimulates the vestibular apparatus, the group of organs in the inner ear responsible for monitoring a person's own balance and movement, then this will induce a sense of self-motion (Todd 1999). But a perceptual approach allows for the experience of either self-motion or the motion of other objects. The relativity of motion ("am I moving relative to the surroundings, or are the surroundings moving relative to me?") means that there is always potentially an uncertainty: when you look out of the window of a stationary train and see movement in relation to another train, the well-known perceptual jolt, as the apparent self-motion is suddenly revealed to be actual movement of the other train, is an illustration of this relativity. In a similar way, sound specifies motion but on its own cannot unambiguously specify the relativities of that motion: in terms of pitch shift, the Doppler effect, for example, is identical whether caused by a sound-emitting object approaching and passing a stationary observer, or a moving observer passing a stationary sound-emitting object. There is little empirical evidence about the extent to which listeners perceive musical motion as self-motion or as the movement of other objects, but some of the language that people use in relation to their musical experiences is indicative of self-motion: "moved," "swept away," "transported," "blown away" are just a few of the terms people use to describe intense musical experiences. In part this may be attributed to a simple principle of ecological acoustics: if all the separate sources (real or virtual) that are specified in a piece of music are heard to move together in a correlated **(p.76)** fashion, this specifies a listener moving in relation to a collection of stationary sound sources (i.e. self-motion). If, however, the various sound sources all move relative to one another, and in relation to the listener, this specifies the movements of external objects in relation to one another. In very simple terms this suggests, for instance, that music with complex polyphonic properties is likely to be heard in the latter category—as the movement of external objects/agents in relation to one another and the listener; while monodic or homophonic music may more easily specify self-motion—movement of the listener in relation to the environment.

Three Musical Examples

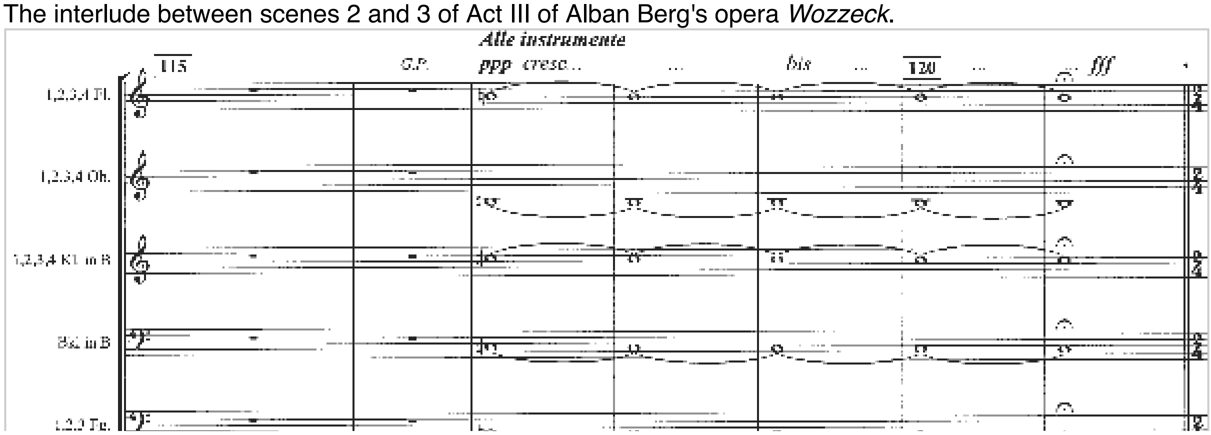
Having considered the perceptual principles that may explain how music specifies motion, this section considers the particular musical features that specify motion in three rather different musical examples. The first example focuses on the pair of dramatic orchestral crescendos on a single pitch (B) that occurs in the interlude between scenes 2 and 3 of the third act of Alban Berg's opera *Wozzeck* (see **example 3.1**). The motion that is specified by these sounds is paradoxical. On the one hand the complete absence of pitch change specifies stasis, while on the other the continuous change in both timbre (in the first of the two crescendos) and dynamic (in both of them) specifies continuous and unidirectional motion. The net result is, perhaps, a sense of highly focused and unswerving approach—the auditory equivalent of what Gibson called “looming” or “time-to-contact.” Gibson writes of the visual information that specifies the approach to, or approach of, an object as follows:

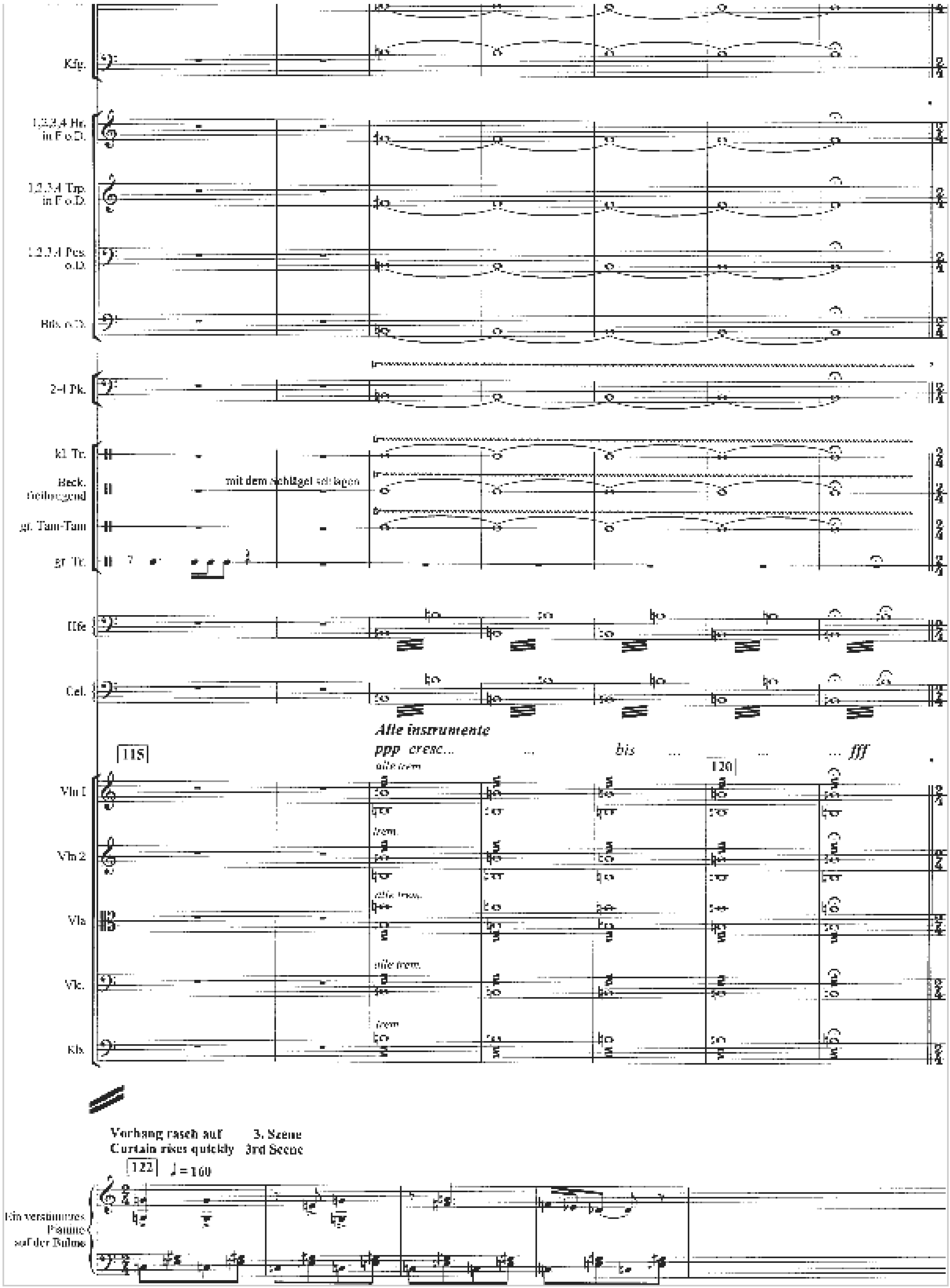
Approach to a solid surface is specified by a centrifugal flow of the texture of the optic array. Approach to an object is specified by a magnification of the closed contour in the array corresponding to the edges of the object. A *uniform* **(p.77)** rate of approach is accompanied by an *accelerated* rate of magnification…. The magnification reaches an explosive rate in the last moments before contact. This accelerated expansion … specifies imminent collision. (Gibson **1958** cited in Gibson **1979** : 231; emphasis in original)

This description of the visual-information specifying approach closely parallels the sounds of these two orchestral crescendos, if one makes appropriate sensory substitutions: continuous dynamic increase substitutes for flow of optical texture, and the pitch stasis provides the centrifugal quality. (Imagine how different the effect of the example would be if the music were to trace some kind of continuous or stepwise pitch trajectory at the same time as the crescendo.) The “explosive rate” of magnification mentioned by Gibson, and the imminent collision that it specifies, is an aspect of the *Wozzeck* example that is largely in the hands of the conductor who controls by exactly how much, and at what rate, Berg's dynamic markings should be realized. Karl Böhm's 1965 recording of the opera, [9] however, does seem to reach this “explosive rate” of intensity increase, and thus the sense of imminent impact.

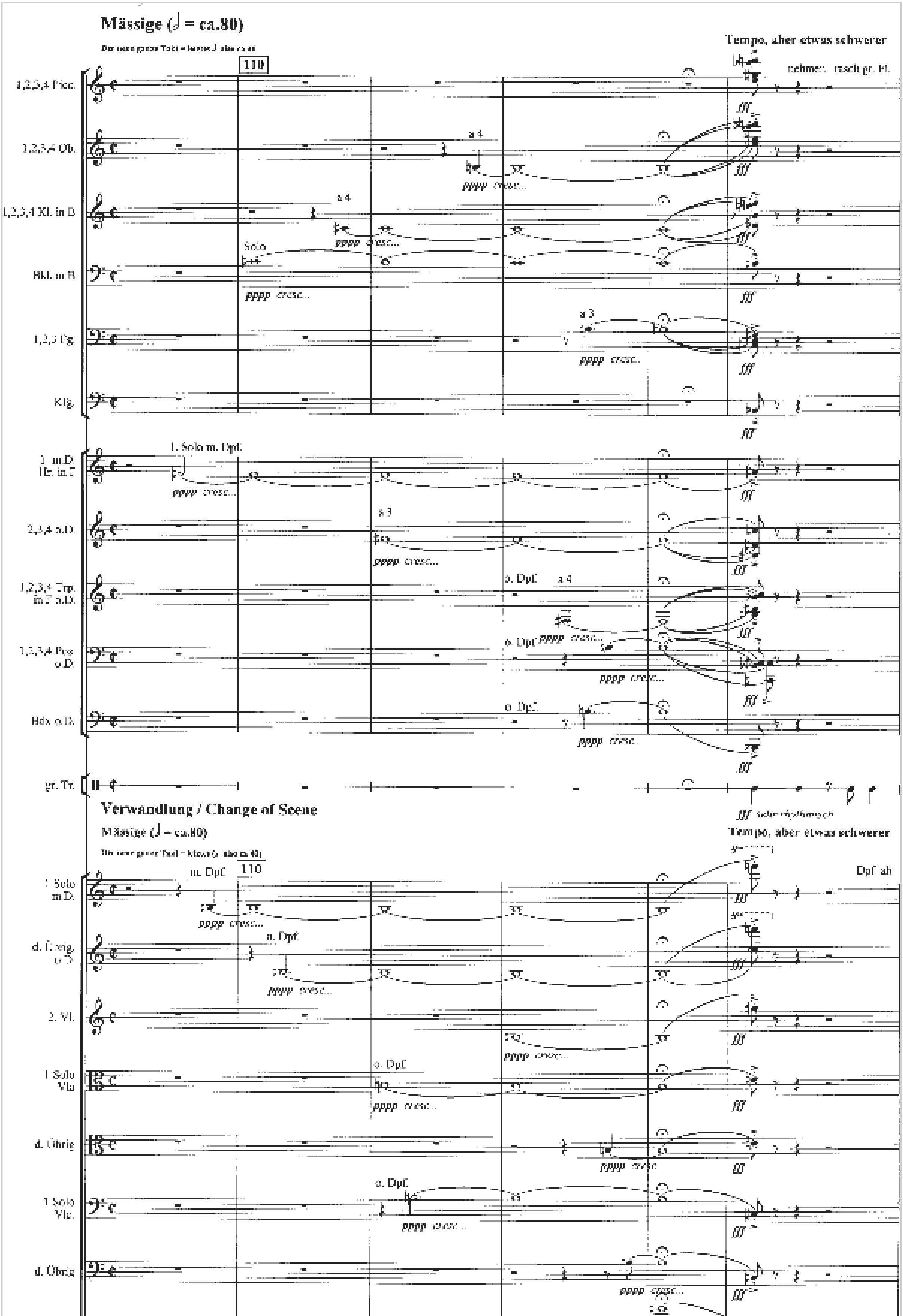
The two crescendos differ in a number of respects and give rise to different motion effects as a result. The first crescendo is achieved not only by continuous increases of dynamic within each of the instruments or instrumental groups of the orchestra, but also by a complex pattern of successive instrumental entries, so that both the timbre and the texture of the orchestral sound change continuously as the dynamic increases. Arguably, this results in a less focused and static “looming” quality than is achieved by the second crescendo, which consists simply of a huge crescendo on an orchestral *tutti*. A second difference concerns the respective endpoints or “contact moments” of the two crescendos. The first, after a build up on unison B, ends with a six-note orchestral chord, which is played as a rhythmic unison on a downbeat, and coincides with the first note of the distinctive rhythmic figure played solo and triple **(p.78)**

The interlude between scenes 2 and 3 of Act III of Alban Berg's opera *Wozzeck*.




Kfg.
1,2,3,4 Hr. in F o.D.
1,2,3,4 Trp. in F o.D.
1,2,3,4 Pos. o.D.
Bts. o.D.
2-4 Pk.
kl. Tr.
Beck. freiliegend
gr. Tam-Tam
gr. Tr.
mit dem Schlägel schlagen
Hfe
Cel.
Alle Instrumente
ppp cresc... ... bis ... ... fff
115
120
Vln I
Vln 2
Vla
Vlc.
Kb.
Vorhang rasch auf 3. Szene
Curtain rises quickly 3rd Scene
122 ♩= 160
Ein verstimmtes Pianino auf der Bühne

(p.79)


Mässige (♩ = ca.80)
110
Tempo, aber etwas schwerer
1,2,3,4 Picc.
1,2,3,4 Ob.
1,2,3,4 Kl. in B
Bkl. in B
1,2,3 Fg.
Kfg.
1. Solo m. Dpf.
2,3,4 o.D.
1,2,3,4 Trp. in F o.D.
1,2,3,4 Pos. o.D.
Btb. o.D.
gr. Tr.
Verwandlung / Change of Scene
Mässige (♩ = ca.80)
1 Solo m.D.
m. Dpf.
2. Vl.
o. Dpf.
1 Solo Vla
d. Übrig
1 Solo Vlc.
d. Übrig
Kb.
pppp cresc...
fff
a 4
a 3
Solo
Dpf. ab

(p.80) forte by the bass drum (see **example 3.1**). The orchestral downbeat has the attack and unanimity of a physical impact, and the sense of motion that the first crescendo conveys is therefore of approach, followed by collision—out of which the bass drum appears with the suddenness of an explosive rupture. The second crescendo, by contrast, is an orchestral unison throughout and consists solely of a dynamic crescendo which ends *not* with a downbeat—indeed without any final "event" at all—but with the equivalent of a cinematic cut straight into the next scene of the opera. (Berg's instruction in the score is that the instruments should be quickly damped, to coincide with the curtain being rapidly raised on the next scene.) It is as if the imminent collision never materializes, and the listener (or the music?) is shot out into a new and completely unexpected space—as if passing through an invisible barrier at the moment when collision seemed inevitable. Using the similarity with optical flow, texture gradients, coordinated versus independent component behavior, and so on (the basis for which is already established within auditory scene analysis research [10]), it would be relatively straightforward to establish empirically what the musical/acoustical conditions are that specify collision, rupture, emergence, unidirectional movement, or movement with frequent directional change, and so on. It simply has not yet been done.

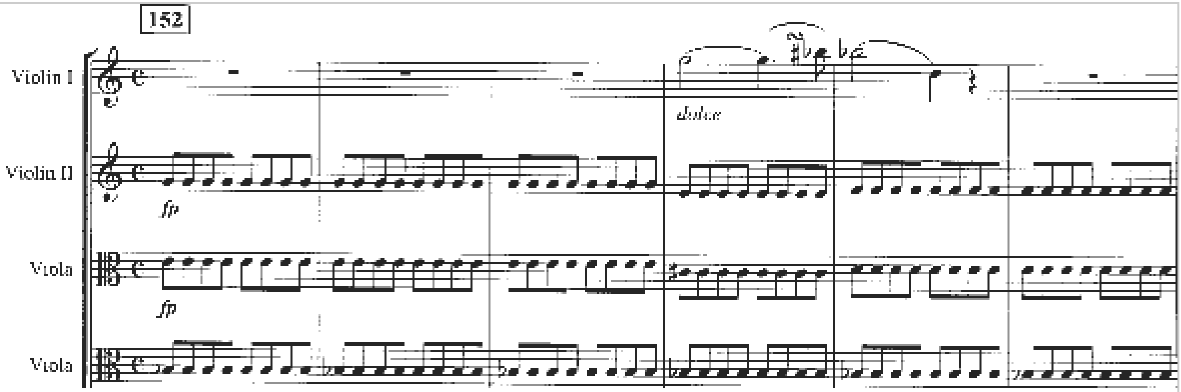
My second example comes from pop music—a short section from the track "Build It Up, Tear It Down" (from the CD *You've Come a Long Way, Baby*) by the British musician Fatboy Slim (Fatboy Slim 1998). The track, which is representative of techniques used in a considerable amount of recent dance music, starts with drum kit and other percussion sounds presented in a relatively dry (nonreverberant) acoustic, soon followed (at 0:15) by a male voice with rather more reverberation. This part of the track demonstrates no particularly unusual motion characteristics. However, at 0:28, the sound and texture abruptly change, a bass instrument of some kind replacing the previous percussion and vocals, playing the pitches B and E in alternation and heard as if through a low-pass filter (i.e. with most of the higher frequency components **(p.81)** eliminated). Over the course of the next half-minute this sound continuously evolves as if the cut-off frequency of the low-pass filter is being progressively raised, the higher frequency components becoming more audible, and revealing the sounds of a drum kit (at about 0:40) and other instruments in the mix. As the filter cut-off rises, resulting in an increasingly bright and clear sound, the same male voice that was heard in the earlier part of the track is added to the mix (at 0:55), reaching a climax of volume, rhythmic density, and timbral brightness until it cuts into another louder and fuller texture (at 1:02) which is the principal material for the rest of the track.

The perceptual effect of the 34-second section described here (from 0:28 to 1:02) is of a continuous movement towards a sound source that is at first occluded but progressively reveals itself. It is not clear whether this specifies self-motion towards a stationary sound source, or a moving sound source that approaches a stationary listening point—but in either case there is a powerful sense of the source being at first concealed below some kind of acoustic horizon, above which it then progressively rises until it is fully revealed and directly in front of the listener by the time the texture changes at 1:02. The explanation for this strong sense of motion, and for the particular *style* of motion described here, comes straight from ecological principles: in everyday circumstances, the acoustic array from a sound source that is concealed in relation to a listener (behind a large object like a wall or a building, or below a horizon) will possess attenuated high frequencies due to simple masking principles. High frequencies are absorbed and dissipated in the environment more rapidly than low frequencies, leading to the characteristic "bass heavy" quality of amplified music heard at a distance (the "open air pop festival" sound). As the distance to the source decreases, or the degree of occlusion declines, the high frequencies increase in relative intensity, shifting the timbral balance towards increased brightness. It is for this reason that the continuous change in filtering on the Fatboy Slim track gives rise to a powerful sense of approach to (or approach by) **(p.82)** a sound source—as does the *Wozzeck* example, but using a different acoustic parameter.

My final example comes from instrumental chamber music and addresses a somewhat different consideration: the role of motion in bringing about a sense of agency in music—of the listener being at times an "overhearer" of musical events, and at others a participant among them. The development section (bars 152–205) of Mozart's String Quintet in C Major (K. 515) features a dramatic shift from one instrumental texture to another and then back again at the start of the recapitulation (see **example 3.2**). The two textures specify very different kinds of perceived motion—differing not in speed, proximity, or trajectory, but in terms of multiplicity and unity. The opening texture of the development section (which is the same as the prevailing texture of much of the exposition) can be described as an interchange between the cello and first violin, with the second violin and two violas supplying a middle-register harmonic “filler” in continuous quavers. The cello and first violin play strongly contrasting material: clearly articulated and linearly rising arpeggios in the cello, more lyrical material in the violin. At bar 168, however, this texture changes abruptly into something far more abstractly contrapuntal, involving the ensemble as a whole, with four of the five instruments at any one time playing overlapping semibreves and minims in a sequence of suspensions and resolutions. One instrumental part always remains more mobile (a neighbor-note figure in quavers which threads its way from the cello up through the ensemble over the course of 15 bars), and from 185 this material becomes increasingly prominent. At 193 recognizable thematic material from the exposition returns, and the clearly differentiated instrumental texture from the start of the development is resumed at the point of recapitulation (at 205).

I hear this passage as embodying a distinct change in perceived motion (and then a change back again)—and one in which there is an associated transformation of the sense of agency. As revealed by the earlier discussion of the relativity of motion and self-motion, the perception of motion in music brings a listener into a very **(p.83)**

Bars 152–205 of the First Movement of Mozart's String Quintet in C Major, K. 515.



(p.84)

(*Continued*)

(p.85)

(p.86) direct relationship with musical materials—as an agent among, or observing, other agents. I hear the opening of this Mozart development section more as an observer than a participant: the cello and violin parts specify very different kinds of motion (one assertive and energetic, the other receptive and accommodating), and for this reason, as well as for reasons of register and timbre, they specify distinct "agents" in motion relative to one another—two separate individuals. I *could* identify with either of them, or perhaps both in a kind of schizophrenic alternation—but actually I experience this as the actions of two separate agents of motion, with myself as onlooker (or "overhearer"). The inner parts have no particularly differentiated motion quality and provide a kind of environment—a ground against which the figures move. The textural change at 168 alters everything: the ground drops away as the inner instruments abruptly cease their regular quaver activity, and the separate parts merge into a single body with a complex movement, one with which I now identify, rather than simply overhear. There is a paradoxical quality to this single body in that its unified character (originating in its contrapuntal unity—a single source) coexists with an obvious sense of independent voices, emphasized by the overlapping, and frequently suspended, nature of the part writing. My own experience of the passage partially resolves the paradox by hearing it as the slow flexing, or intertwining, of the limbs of a single "body" with which I identify. This is, in other words, an experience of *self*-motion—while the earlier part of the development seemed more like the independent movements of others. As the texture again becomes more motivically differentiated (from about 184—see **example 3.2**), this sense of a single body starts to break up into distinct streams again, and by the start of the recapitulation the sense of separate agents, in relation to which I am an overhearer, is regained. The kinds of transformation described here, somewhat reminiscent of the kind of visual morphing that can be accomplished with computer animation, demonstrate how music has the capacity to call into question the **(p.87)** rigid separation of subject and object and to play with states which do not wholly conform with either the one or the other. It shares this characteristic with other forms of "virtual reality": the player in a skateboarding computer game, for example, is commonly both subject (he or she is controlling the board, and the environment is frequently visualized from the perspective of the board) and object (the skateboard and rider will appear as an object in the scene) in a manner that the users of such videogames find intuitively convincing and unproblematic.

Another issue that these examples broach is the question of who or what is moving. As all three analyses illustrate, there is an ambiguity about the agency to which the movements described above should be attributed. In the case of *Wozzeck*, this is made more concrete and particular by the operatic context of the music and the drama of which it is a part. Does a listener hear him or herself as moving towards some collision, or as one of the characters of the opera moving towards some collision, or indeed as some other person or object in the drama? Given the dramatic context in which this brief musical interlude occurs (the main character, Wozzeck, has just murdered his lover, Marie, and will shortly drown himself), listeners may hear this as Wozzeck "rushing to meet his fate" [11] (or fate rushing to overtake him), or perhaps death rushing to meet Marie (heard now in immediate retrospect). But it is equally possible for a listener to hear this as self-motion, through an identification with one of the opera's characters—for me to hear myself as Wozzeck, for instance, hurtling towards disaster.

A reader might object that this discussion of the example from *Wozzeck* is unacceptably speculative, full of interpretive license, and basically out of step with an ecological approach: it seems to depend heavily on verbal and dramatic information to interpret the perceptual information (the sounds of the two orchestral crescendos)—rather than relying on specification by stimulus invariants. But this overlooks the fact that all of the elements mentioned (the **(p.88)** drama, the characters, the sounds) *are* part of the available information for a viewer/listener. The considerable leeway that exists for different perceptions of this short musical extract is not a problem for an ecological approach: many much more mundane environmental circumstances can be perceived in more than one way, and aesthetic objects are particularly multivalent. Not only do the objects themselves often contain deliberately partial or conflicting perceptual information, with the consequence that they afford multiple possibilities; but the viewers/listeners who encounter them, even when drawn from notionally the same culture, may also differ markedly in their previous experience of these or similar events. Musical sounds are "underdetermined": on their own, they may often seem to specify many things—or maybe nothing in particular. But this can be strongly affected by the wider context in which they are heard: the regular quaver chords of the piano part in “Gute Nacht,” the first song of Schubert's song cycle *Winterreise*, can be heard as specifying all kinds of motions and scenes, but a listener who knows the cycle's title, or has read the text of the first song in a program note, or even just seen a CD cover with an illustration of a wintry scene, is likely to hear this as footsteps in the snow; the information in the text or image serves to focus dramatically the potentially much wider range of scenes and actions that these sounds can specify. In this sense, as Nicholas Cook has argued (Cook **1998** a; **2002**), music almost always has a multimedia quality to it, and musical meaning is always the consequence of a context that is wider than the “sounds in themselves.” The Fatboy Slim and Mozart examples leave open the interpretation of the scenes and motions that the sounds specify, as also does the *Wozzeck* example, even with its much more programmatic textual and dramatic context. But the general ecological principles proposed here constitute an explanatory framework that helps to understand not only the constraints that apply in this process, but also why there should be any perception of motion and agency in music in the first place.

(p.89) Summary and Implications

This chapter has proposed that the sense of motion and gesture in music is a straightforwardly perceptual phenomenon, and that the process by which auditory information specifies motion in music is, broadly speaking, the same as the process by which motion is specified in everyday circumstances. Part of the motion that listeners perceive may be the real movements of the performers and instruments involved (Shove and Repp **1995**), but an important component of the motion in music is neither real nor metaphorical, but fictional —in the same way that the scene portrayed in a picture may be fictional. The sense of motion or self-motion draws a listener into an engagement with the musical materials in a particularly dynamic manner (he or she seems to act among the materials), and in doing so constitutes a vital part of musical meaning. In particular, the sense of motion or self-motion raises intriguing questions of agency, as I have tried to show with the three musical examples presented here. Who or what is moving, with what style of movement, to what purpose (if any), and in what kind of virtual space? The theory proposed here has specific empirical implications that could be pursued—most obviously an investigation of the various kinds of perceptual information in music that specify motion and a consideration of whether they function in the same manner as perceptual information does for the motion of objects in everyday circumstances.

A number of authors have dealt with questions of subjectivity in music from various perspectives (e.g. Cumming **2000** ; DeNora **2000** ; Lidov **1987**), and Watt and Ash (**1998**) have carried out empirical work demonstrating that listeners identify “person-like” qualities in music more readily than other types of attributes. Starting from a different perspective, but converging with the ideas presented here, Watt and Ash (**1998** : 49) conclude that “loosely speaking, music creates a virtual person.” A common feature in all of this writing is the recognition that a listener's sense of meaning in **(p.90)** music is powerfully bound up with his or her experience of being subjectively engaged (or alienated) by the music, and with the varieties of subjective states that music can afford. An important component of that subjective engagement with music is its corporeal, proprioceptive, and motional quality, which may on occasion provide listeners with experiences of “impossible worlds” that have some of the same attractions as do other forms of virtual reality. [12] Just as Cook (**1998** b) has argued that music is a means of gaining insight into other cultures and histories, and that we listen to music “not just for the good sounds, though there is that, but in order to gain some insight into those (sub)cultures” (129), so too music affords peculiarly direct insight into a limitless variety of subjective experiences of motion and embodiment—real and virtual.