

SEMPRE Studies in The Psychology of Music

Series Editors

Graham Welch, *Institute of Education, University of London, UK*
Adam Ockelford, *Roehampton University, UK*
Ian Cross, *University of Cambridge, UK*

The theme for the series is the psychology of music, broadly defined. Topics will include: (i) musical development at different ages, (ii) exceptional musical development in the context of special educational needs, (iii) musical cognition and context, (iv) culture, mind and music, (v) micro to macro perspectives on the impact of music on the individual (such as from neurological studies through to social psychology), (vi) the development of advanced performance skills and (vii) affective perspectives on musical learning. The series will present the implications of research findings for a wide readership, including user-groups (music teachers, policy makers, parents), as well as the international academic and research communities. The distinguishing features of the series will be this broad focus (drawing on basic and applied research from across the globe) under the umbrella of SEMPRE's distinctive mission, which is to promote and ensure coherent and symbiotic links between education, music and psychology research.

Other titles in the series

Sociology and Music Education
Edited by Ruth Wright

The Musical Ear: Oral Tradition in the USA
Anne Dhu McLucas

Infant Musicality
New Research for Educators and Parents
Johannella Tafuri

New Perspectives on Music and Gesture

Edited by

ANTHONY GRITTEN
Middlesex University, UK

and

ELAINE KING
University of Hull, UK

ASHGATE

- Ricoeur, P. (1981). *Hermeneutics and the Human Sciences*. Cambridge: Cambridge University Press.
- Rosenbaum, D. A. (1991). *Human Motor Control*. San Diego: Academic Press.
- Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J. & van der Wel, R. (2007). The Problem of Serial Order in Behavior: Lashley's Legacy. *Human Movement Science* 26: 525–54.
- Schaeffer, P. (1966). *Traité des Objets Musicaux*. Paris: Éditions du Seuil.
- (1998). *Solfège de l'Objet Sonore*. Paris: INA/GRM.
- Schleidt, M., & Kien, J. (1997). Segmentation in Behavior and What it can Tell us about Brain Function. *Human Nature* 8: 77–111.
- Schoenberg, A. (1967). *Fundamentals of Musical Composition*. London: Faber & Faber.
- Snyder, B. (2000). *Music and Memory: An Introduction*. Cambridge, MA: MIT Press.
- University of Oslo (2010). fourMs - Music, Mind, Motion, Machines (<http://fourms.uio.no>; accessed 18 November 2010).
- Varela, F. (1999). The Specious Present: The Neurophenomenology of Time Consciousness. In J. Petitot, F. J. Varela, B. Pachoud & J. M. Roy (eds), *Naturalizing Phenomenology* (pp. 266–314). Stanford: Stanford University Press.
- Vines, B., Dalca, I. & Wanderley, M. (2006). Variation in Expressive Physical Gestures of Clarinetists. In M. Baroni, A. R. Addressi, R. Caterina & M. Costa (eds), *Proceedings of the Ninth International Conference on Music Perception & Cognition (ICMPC9), Bologna, 22–26 August 2006* (pp. 1721–2). Bologna: SMPC & ESCOM.
- Wiesendanger, M., Baader, A. & Kazennikov, O. (2006). Fingering and Bowing in Violinists: A Motor Control Approach. In E. Altenmüller, M. Wiesendanger & J. Kesselring (eds), *Music, Motor Control and the Brain* (pp. 109–23). Oxford: Oxford University Press.
- Wohlschläger, A., Gattis, M. & Bekkering, H. (2003). Action Generation and Action Perception in Imitation: An Instance of the Ideomotor Principle. *Philosophical Transactions of the Royal Society of London B* 358: 501–15.
- Woodworth, R. S. (1899). The Accuracy of Voluntary Movement. *Psychological Review* 3: 1–119.

Chapter 4

Musical Gesture and Musical Grammar:
A Cognitive Approach

Lawrence M. Zbikowski

A case could be made that our notion of musical gesture is a thoroughly metaphorical one. Physical gestures – pointing with one's hand, or nodding one's head – are typically soundless. Musical gestures, by contrast, are all about sound, or more specifically about *sequences* of sounds. And although musical gestures are often (although not exclusively) produced by physical actions, there may be no correlation between the gesture and the sound that produced it: an arch-like musical gesture, for instance, might well be produced by a decidedly lateral motion with no appreciable up and down component. Finally, physical gestures – and especially the sort of physical gestures on which I would like to focus – are often correlated with speech, as when the phrase 'The book is over there' is accompanied by the speaker pointing with their hand, or when the utterance 'That's a quite good idea' is accompanied by nodding. In both cases the gesture adds something crucial to communication: in the first case, an indication of the location of the book (that is, a specification of where 'there' is); in the second case, adding affirmation to the evaluation (the idea is not only a good one, but one with which the speaker agrees). To expand on this point just a bit, consider how our understanding would change if the statement 'That's a quite good idea' were accompanied by the speaker *shaking* their head. One interpretation of such behaviour might be that the speaker is both sharing their positive evaluation of the idea and wondering why they didn't think of it first. In all three cases the physical gestures add information that is not included in the linguistic constructions uttered by the speaker. The same, generally speaking, is not true of musical gestures – or at least is not true of musical gestures as we usually think of them.

On this evidence, it would seem that the notion of musical gesture is indeed a metaphorical one. We can add some particularity to this assessment through the framework provided by the contemporary theory of metaphor (Grady 2007; Lakoff 1993; Zbikowski 2008b): 'musical gesture', as it is conventionally construed, reflects a conceptual mapping in which knowledge from one domain of experience (namely, *physical gestures*) is used to structure another domain of experience (*sequences of musical materials*) with the goal of organizing our understanding of the second domain. According to this interpretation, when we describe the sequence of pitches E–F#–G–F#–E as an 'arch-like gesture' we are using our knowledge about physical gestures (specifically, those that trace the

pattern of an arch) to characterize musical sounds. This mapping offers a visual analogue for the musical sounds (and here I would wish to set to one side the way the pitches are rendered in staff notation, which I regard as supporting but not necessary to their visualization) as well as a kinesthetic analogue: the pitches of the musical gesture 'rise and fall' (as does a physical gesture that traces an arch) and there is a sense of repose when, after this trajectory, the sequence of pitches returns to its starting point.

While I am perfectly comfortable with this perspective on musical gesture, and would agree that many descriptions of musical gesture are based on this sort of metaphorical mapping, in what follows I would like to explore the idea that the communicative function demonstrated by the physical gestures that accompany speech is one shared by sequences of musical materials. Here I should be a bit more specific about the physical gestures on which I wish to focus, which are somewhat different from deictic gestures such as pointing, or conventional gestures such as nodding. The gestures that are my concern are more spontaneous, created on the fly in the process of discourse. In one attested case, for instance, a speaker accompanied the utterance 'And he went up a winding staircase' by making a spiral motion with his index finger pointing upward, a gesture that made clear that the ascent was up a circular staircase (Iverson and Goldin-Meadow 1997: 458). Again, the information provided by the gesture is distinct from – and even of a different sort than – that provided by the utterance: where the linguistic construction gives us a general picture of the events (a person ascending a staircase) the gesture gives us a dynamic analogue for the actual process of ascent, from which we then infer the physical characteristics of the staircase. I would like to propose that the sequences of sound events proper to music can, in a similar fashion, provide analogues for dynamic processes. There are, however, two important differences between these two means of creating analogues for dynamic processes. First, where the physical gestures that accompany speech are closely linked to language, the sonic analogues of music are largely independent of language. Second, while it is generally recognized that the gestures that accompany speech do not have a grammar, I would like to propose that sonic analogues are *basic* to musical grammar. From this perspective, then, the notion of a musical gesture is not a metaphorical one but a reflection of the essential materials of musical expression.

I would like to use this chapter to explore some of these topics in a bit more detail. In the first section I shall sketch a cognitively based approach to musical structure, using as my example Jerome Kern and Dorothy Fields's 'The Way You Look Tonight', a tune written for the 1935 Fred Astaire and Ginger Rogers' vehicle *Swing Time*. In the second section, I shall review recent work by the psycholinguist David McNeill and others on the spontaneous gestures that accompany speech, with particular attention to how the thought processes associated with such gestures correlate with the thought processes revealed through speech. McNeill's analysis of the structure of gestures can help us understand some of the features of 'The Way You Look Tonight', a connection I explore in the third section, in which I shall also elaborate further my approach to musical grammar. In the fourth section, I want

to take a closer look at the gestures Fred Astaire makes as he performs the song in the course of the movie, gestures that suggest a very close relationship between the basic materials of music and the physical gestures that accompany language, and that can help us better understand some of the properties of musical grammar.

The Musical Materials of 'The Way You Look Tonight'

Let me begin, then, with an analysis of some of the musical materials Kern made use of in crafting 'The Way You Look Tonight' (Example 4.1 provides a lead sheet for the song). As do many of the songs of this period, 'The Way You Look Tonight' has an AABA' form; in what follows my principal focus will be on the melody of the first A section. One way to both describe and hear this melody is in terms of a number of smaller units. There is, for instance, the falling fifth of bars 1 and 2, which is answered by the arch-like rise and fall through a third that occupies bars 3 and 4. The latter motion is replicated (a step higher) in bars 5 and 6, and again in bars 7 and 8. This last arch turns out to be a bit more complicated, for as soon as it has returned to its starting G it ascends once more, this time continuing up to the D of bar 9. After the continuous crotchets of bars 7 and 8 there is a leisurely fall from the D of bar 9 to the D of bar 10, a descent through an octave that both echoes and expands the descent through a fifth of bars 1 and 2. (The connection between these two moments in the melody is all the stronger in that they both stand at the beginning of an eight-bar sub-phrase, and both conclude on D.) The octave descent of bars 9 and 10 is followed by a figure that ascends and descends, much as did the figures of bars 3 and 4, and 5 and 6, but with important differences: the arch is not as smooth; the durations of the notes become longer (in successive stages); and the entire figure ends on D, a whole step below where it started.

The small units that I have described – the opening falling fifth, or the arch-like melody of bars 3 and 4 – are often called musical motives; in that these units combine pitch and rhythmic materials, they are also similar to Arnold Schoenberg's *Grundgestalten*. As part of my efforts to approach such basic materials in a way that conforms with recent research in cognitive science, I have suggested that such units could be thought of as categories of musical events (Zbikowski 2002: Chapter 1). Bars 1 and 2, and 9 and 10, for instance, could be put into one category – let's call it Category X – and bars 3 and 4, and 5 and 6, could be put into another – Category Y. Note that all of the members of a category need not be identical to one another. For instance, bars 7 and 8 could well be included in Category Y, but with the understanding that they are less typical of the category than are bars 3 and 4, or 5 and 6. In general, the categories basic to musical understanding demonstrate what are called typicality effects: membership in the category is graded, and some members are more typical of the category than are others.

Although it may seem somewhat cumbersome to think of the constituent pitch sequences of 'The Way You Look Tonight' as members of a cognitive category rather than musical motives or instances of a *Grundgestalt*, this approach comes

Example 4.1 Lead sheet for 'The Way You Look Tonight' by Jerome Kern and Dorothy Fields

SOME - DAY WHEN I'M ALL - FLY LOW, WHEN THE WORLD IS COLD, I WILL FEEL A GLOW JUST THINK - ING OF YOU.
 AND THE WAY YOU LOOK TO - NIGHT OH, BUT YOU'RE LOVE - LY WITH YOUR SMILE SO WARM.
 AND YOUR CHEEK SO SOFT, THERE IS NOTH - ING FOR ME BUT TO LOVE YOU, JUST THE WAY YOU LOOK TO - NIGHT.
 WITH EACH WORD YOUR TEN - DER - NESS GROWS, TEAR - ING MY FEEL - A - PART.
 AND THAT LAUGH THAT WREN - BLES YOUR NOSE, TOUCH - ES MY FOOL - ISH HEART LOVE - LY.
 HE - VEE, HE - VEE CHANGE, KEEP THAT BREATH - LESS GUSH, WON'T YOU PLEASE AT - TEND ME I LOVE YOU, JUST THE WAY YOU
 LOOK TO - NIGHT MM - MM - MM - MM JUST THE WAY YOU LOOK TO - NIGHT

with two advantages. First, by drawing on the explanatory framework provided by empirical research on typicality effects in categories it becomes possible to describe in a principled and consistent way relationships between non-identical but similar sequences of musical materials. Although I shall not pursue such a description here (having already done so for works by Mozart and Beethoven in Zbikowski 1999 and 2002), this methodology offers an important resource to those who would seek to understand relationships between musical materials – and not just those conventionally described as ‘motives’ – over the course of a work of music. Second, and more significant for the perspective I wish to develop in this chapter, thinking of the basic materials of ‘The Way You Look Tonight’ in terms of cognitive categories connects these materials with other sorts of categories through which we structure our understanding of the world. This connection can help us to come to terms with the part music plays in our cognitive lives, both as an object of understanding and as a means of structuring knowledge.

With these thoughts in mind, I would like to propose that categorization – understood as an active cognitive process, one used by humans to structure their understanding of the world – offers a way to explain how musical communication comes about. On the one hand, listeners actively organize sound phenomena into musical categories to make sense of what they hear. On the other hand,

composers and improvisers structure sequences of musical events so that listeners have easy access to elements out of which musical categories can be built. The overarching idea behind this perspective is that human cognitive capacities shape the production and reception of music. In the approach to musical structure that I have developed over the past few years, I have been interested in how cognitive capacities associated with categorization, as well as those associated with analogy and the use of conceptual models, are manifested in our understanding of music, taking the approach that these capacities inform both the organization and the meaning of music. As I view it, both the organization and meaning of music are central to its grammar, a notion I shall return to toward the end of this chapter.

Gesture and Thought

An interest in the cognitive resources behind communication of the sort that has informed my research on musical structure can also be seen in research on gesture and language. Over the past 25 years David McNeill and his colleagues have developed compelling evidence that the gesture that accompanies language introduces new information into discourse structures. Put another way, the gestures that accompany our speech reflect a mode of thought that is independent from but coordinated with language. Understanding why some parts of our thought are captured by gesture and why others are captured by language gives crucial insight into both the goals of human communication and the way these goals are accomplished.

One way McNeill studied the relationship between gesture and language was to show subjects a short, action-packed, but wordless clip from a cartoon and then ask them to describe what they saw. In a typical example, a 40-second long clip showed a scene in which Sylvester the cat attempts to get to Tweety Bird, who is apparently safe and sound in an apartment many stories above the street, by climbing up the inside of a drainpipe. Tweety manages to thwart this effort by retrieving a bowling ball and dropping it down the drainpipe. The viewer sees the progress of the bowling ball down the drainpipe as well as the moment when it comes into forcible contact with Sylvester, rendered as an almighty crash and noticeable increase in the bulge of the drainpipe. The bowling ball then continues its descent, accompanied by Sylvester, and when both emerge from the end of the drainpipe we see that Sylvester has in fact swallowed the bowling ball. The bowling ball, now inside Sylvester, begins to roll down the street, gradually gathering speed until, at the bottom of the street, it enters a bowling alley, with Sylvester helplessly pulled along the whole way. A moment later we hear the crash of the bowling ball hitting the bowling pins, the force and volume of which suggests that Sylvester has achieved a strike.

In his discussion of these experiments McNeill provided a relatively full transcript of the responses of one subject identified as Viv., which is provided in Table 4.1. In the videotape from which the transcript was drawn (which runs

about 20 seconds), Viv. sits in a chair, addressing an off-camera interviewer. As she speaks, she makes a series of small, relatively contained gestures with her hands. McNeill divides Viv.'s description up in accordance with the gestures she makes, yielding the nine lines indicated in the leftmost column of the table; square brackets are used to indicate the initiation and conclusion of each gesture stroke. (The gesture stroke in line 2, for instance, does not begin until Viv. has begun to say 'ball'.)

Table 4.1 McNeill's analysis of gestural catchments in Viv.'s description of the Sylvester and Tweety Bird scene (adapted from McNeill 2005: 118)

Line	Utterance	Gesture feature	Catchment
1	He tries going [up the inside of the drainpipe and]	1-hand (right)	C1
2	Tweety Bird runs and gets a bowling ball and drops it down the drainpipe]	2-similar hands	C2
3	[and as he's coming up]	2-different hands	C3
4	[and the bowling ball is coming d]	2-different hands	C3
5	[own he swallows it]	2-different hands	C3
6	[and he comes out the bottom of the drain]	1-hand (left)	C1, C3
7	[npipe and he's got this big bowling ball inside h]im	2-similar hands	C2
8	[and he rolls on down into a bowling ball]	2-similar hands	C2
9	[ey and then you hear a strike]	2-similar hands	C2

In his analysis, McNeill focuses on a small number of gestural features that recur over the course of the description; McNeill calls these structures *catchments*. In general usage, a catchment is something that temporarily catches water – the successive basins of a waterfall are a helpful image. The catchments with which McNeill is concerned are imagistic rather than hydraulic – in his words, 'A catchment is a kind of thread of visuospatial imagery that runs through a discourse to reveal the larger discourse units that encompass the otherwise separate parts' (2005: 116–117). More formally, we might describe a catchment as a distinctive sequence of physical movements that can be combined with similar sequences to create an analogue for a series of events relevant to discourse. McNeill identifies three catchments in this section of Viv.'s discourse. Catchment 1 is concerned with a single moving entity; its recurring gestural feature makes use of just one

hand. Catchment 2 is concerned with the bowling ball; its recurring feature is a rounded shape using two similarly shaped hands. Catchment 3 concerns the relative position of the bowling ball and Sylvester in the drainpipe; its recurring feature involves two hands that, through their different functions, represent the appropriate spatial configuration.

What is remarkable about the gestures and language of Viv.'s description is the insight they give into her view of Tweety Bird's role in the scene. Were one to read only the first two lines of the transcript it would be easy to accord Tweety Bird a fairly significant role – after all, he is the agent that drops the bowling ball and sends Sylvester on his way to a strike. Viv.'s gestures, however, tell a different story: her gesture in line 2 begins not with the mention of Tweety Bird but with the description of the bowling ball dropping. As the description continues, the focus is on the bowling ball as an antagonistic force, one that shapes the rest of the scene. The gesture of line 2 thus marks the beginning of a process through which gesture and language, working together, structure thought. McNeill calls the whole of this process a *growth point*. As he puts it, 'Growth points are inferred from the totality of communicative events, with special focus on speech–gesture synchrony and co-expressivity. Semiotically a combination of opposites, image [provided by gesture] and form [provided by language], the growth point creates a benign instability that fuels thought and speech' (McNeill 2005: 105). In the present case, the relevant instability is one that shifts focus onto the bowling ball, and turns it, rather than Tweety Bird, into the main antagonistic force of the scene that is then described.

McNeill's research demonstrates two important things about gesture and its relationship to language. First, gesture offers a dynamic, imagistic resource for conveying thoughts that would be cumbersome to express through language. Second, gesture informs and shapes our use of language. As McNeill, Adam Kendon (2004), Susan Goldin-Meadow (2003a, 2003b), and a number of other writers have argued (Duncan et al. 2007), to express our thought we need both the dynamic, imagistic resources of gesture and the firm, stable constructs of language.

Musical Gesture, Physical Gesture, and Musical Grammar

The musical categories I described in my analysis of 'The Way You Look Tonight' might well seem like the firm, stable constructs we typically associate with language, but this is really just an artefact of my analysis. In truth, I think of such categories as highly dynamic structures, analogous in many respects to the catchments described by McNeill. As do gestural catchments, musical catchments contribute to the development of something analogous to a growth point. To explore these ideas, let's return to Kern's tune, this time turning our attention to Dorothy Fields's lyrics. I shall again concentrate on the first verse, with only a few comments on the second and third verse.

The first verse of Dorothy Fields's lyrics is framed with respect to a future state of affairs, which is invoked by the opening 'Someday'. At this future point the speaker, beset with rather dire circumstances – depression and an uncaring world – will be comforted by the remembrance of the object of his affections and, more specifically, by the way she looks on this particular night. As is hinted at by the transformative effect of the appearance of the beloved, and confirmed by the second verse, what is involved here is not simply a kind of passive looking, with one person gazing on another, but an intimacy of association that has both power and depth. The trajectory of the first verse is not simply into a future that looks back to the present as a golden past, but one that establishes 'the way you look tonight' as central to a highly charged romantic relationship.

The trajectories described by the verse have their correlate in the organization of the melody of the tune. Bars 1 and 2, and 9 and 10 – that is, the members of Category X – serve to anchor the tonal frame. These bars are correlated with the words that establish the temporal context that frames the lyrics (the 'Someday' of bars 1–2), and that introduce the beloved (the presumed referent for the 'of you' in bars 9–10). Bars 3 to 8, populated by members of Category Y, explore and expand this tonal frame, in part through destabilizing it. Note, for instance, that the D major harmony implied by bar 5 (and confirmed by the chord symbol) does not act as a stable tonic, but as a way station along an ascending sequential path. It is also in these bars that the lyrics fill out the scene ('When I'm awf'ly low / and the world is cold') and set up the motivation for the contemplation of the beloved ('I will feel a glow just thinking [of you]'). Bars 11 to 14 are occupied with a cadential figure that brings the melody to a close, both through a deceleration of the rhythmic activity and a scalewise descent from A to D. These bars, of course, set the lyrics central to the conceptual frame of the verse: 'the way you look tonight'.

Again, I view these categories as analogous to gestural catchments. As do gestural catchments, these musical catchments interact with and shape the story that is told through language. As one example, in the lyrics set in bars 3–6, which focus on things with a negative valence, the music proceeds with materials typically associated with a positive valence – that is, an orderly, sequential major-mode melody that ascends. This strategy redirects the focus of the line away from the immediate thoughts presented in the lyrics for bars 3–6 and toward the ideas presented in bars 7–10. It is these ideas that draw everything together into a coherent whole with a particularity aimed at the beloved: 'I will feel a glow just thinking of you'. The effect is analogous to that of a growth point: ideas presented by the music and by the lyrics come together in a dynamic process that yields a single unified thought. Similar growth points can be seen in the second and third verses. In the second verse, it is the compelling, even bewitching attractiveness of the beloved that emerges: 'There is nothing for me but to love you'. In the third verse, what emerges is the speaker's desire for reciprocity – the beloved is asked to acknowledge the speaker's love by never changing: 'Won't you please arrange it, 'cause I love you'.

In my review of McNeill's approach to the analysis of gestures I suggested that we could think of a gestural catchment as a distinctive sequence of physical movements that can be combined with similar sequences to create an analogue for a series of events relevant to discourse. As I conceive of it, a musical catchment is a distinctive sequence of sound events that can be combined with similar sequences to create an analogue for a dynamic process. In the case of 'The Way You Look Tonight', the sonic analogues of music – and the dynamic processes with which they are concerned – occur on two levels, one local and quite specific, the other more global and rather less specific. Together, these various sonic analogues not only provide an analogue for gestural catchments, but also the basis for the musical grammar of the passage.

The local level of sonic analogues is occupied with the musical materials of Category X, Category Y and the cadential figure. We need not anchor these categories to specific dynamic processes, but can instead simply note how the kinds of movement they summon contrast with one another: Category X as expansive and leisurely, for instance, and Category Y as more compact, with more of a sense of directed movement. On a more global level, these sonic analogues are organized by syntactic processes. The first such process sets out, in bars 1–4, the materials for musical discourse: Category X and Category Y. These materials are then explored and ramified in bars 5–14. The expansion of the tonal frame enacted in bars 5–8 exploits the open-ended nature of the presentation of Categories X and Y (concluding, as it does, with the melody on scale step 2 in bar 4, harmonized with a dominant seventh chord), and makes use of modified repetitions of Category Y to explore the materials and tonal context established in bars 1–4. When a member of Category X returns, in bars 9–10, it serves both to complete this development of material (as an almost inevitable consequence of the ascending scale of bar 8), to arrest it (by replacing the steady crotchets of the preceding two bars with minims) and to break the regular sequence of two-bar sub-phrases in place from the beginning of the tune. This is, in turn, followed by the cadential figure of bars 11–14, which, as mentioned above, brings the melody to a definitive close. The syntactic processes that organize the musical materials of Category X, Category Y and the cadential figure – the first presenting the topics for musical discourse, the second developing those topics and the third closing off the process of development – are themselves analogues for dynamic processes, but made up not of individual sequences of sounds but of *categories* of such sequences.

I will readily acknowledge that this is a rather novel approach to musical grammar, but it is one that conforms, in its broad outlines, with similar approaches developed by cognitive linguists over the past 20 years. One of the basic assumptions shared by these approaches is that form and function – syntax and semantics, if you like – are deeply linked in the grammar of any mode of communication. With this in mind, I would like to suggest that one of the reasons humans have kept both language and music around is because they have different functions. The most primary and basic function of language within human culture is to direct the attention of another person to objects or concepts within a shared referential

frame (Tomasello 1999: Chapter 5). The most primary and basic function of music is to provide sonic analogues for various dynamic processes that are central to human experience, processes that include the movements of our body through space (Zbikowski 2008a), and the physiological transformations associated with emotions. In the grammar of a mode of communication, these functions are linked with forms; this is the basis of what linguists call construction grammar. In the case of language, constructions are 'stored pairings of form and function, including morphemes, words, idioms, partially lexically filled and fully general linguistic patterns' (Goldberg 2003: 219). In the case of music, constructions are sequences of musical materials – such as the members of Category X and Category Y, or the cadential figure that closes off each of the verses – that serve as sonic analogues for dynamic processes. These basic constructions are organized into larger structures through syntactic processes such as those that shaped the melody for the first verse of 'The Way You Look Tonight'; in keeping with the construction grammar approach, in which *all* grammatical elements are pairings of form and function, syntactic processes are themselves sonic analogues for dynamic processes.

It has often been the case that when music theorists have tried to formalize musical grammar they have drawn on their knowledge of linguistic grammar, relying on the sort of metaphorical mapping that has also informed thinking about musical gesture. Thus Heinrich Christoph Koch, writing in 1787, drew comparisons between the organization of language and the organization of music, and suggested that a complete sentence in speech, with subject and predicate, could serve as a model for a complete phrase in music (Koch 1983[1787]: 4). Although I have no doubt that some similarities between music and language obtain, inasmuch as both are the products of human cultures, the different functions each serve within human culture suggests that their grammars will also be different – indeed, it is grammar that makes it possible for each form of communication to realize its function. What these grammars *do* have in common is a reliance on basic features of human cognitive processing, something also shared by the sonic analogues of music and the physical gestures that accompany speech.

Moving to Music

I have been careful, up to this point, to describe musical materials and musical processes as being *analogous* to things that happen in the world of physical gestures. This is not simply technical fastidiousness on my part: in the account of musical organization I seek to develop I want to keep music and gesture separate, if only because music involves sound and gesture typically does not. My assumption, however, is that the two are deeply linked in cognitive organization. Support for this sort of linkage is provided by a scene from the movie *Swing Time*, in which 'The Way You Look Tonight' made its first appearance. In the movie, Fred Astaire plays a dancer who has recently come to New York to earn some money. He encounters Ginger Rogers, who plays a dance instructor, and soon has

designs on her, both as a dance partner and a possible romantic interest. In the relevant scene, Astaire is trying to win Rogers over – he has managed to get in to her hotel suite, and, even though she has repeatedly spurned him, gives it his best effort by singing 'The Way You Look Tonight'. Rogers, for her part, has locked herself in her room and started shampooing her hair, which provides the basis a dramatic situation that is both intimate and comedic.

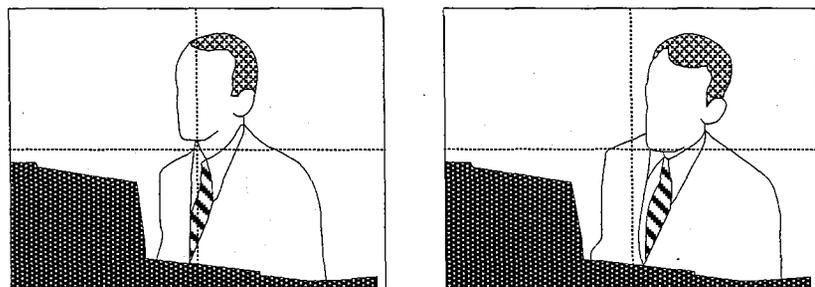
Let me once more focus my comments on the first verse, but in this case because in the movie the performance of the verse is captured in one sustained take, which uses only one camera angle and which has as its sole object Astaire's performance; a still from the opening of the scene is given in Figure 4.1. For my discussion of the gestures Astaire makes in the course of his performance I will rely on line drawings created from still shots taken from the movie; although these drawings will necessarily leave out some very interesting details, they will allow us to focus on the principal features of Astaire's movements.



Figure 4.1 Still of Fred Astaire at the piano from *Swing Time*

On 'Someday', Astaire moves toward the right side of the frame, lifting his shoulders and chin slightly, and then concludes the movement by dipping his head toward the keyboard; the beginning and end of the gesture are shown in Figure 4.2. This gesture is typical of those he makes in his performance: constrained by the conceit that he is accompanying himself on the piano, his movements are restricted to those he can make with his head and his torso. The movements he makes with his arms are much less noticeable and, in that the camera angle takes a view of Astaire from over the top of the piano, are almost invisible.

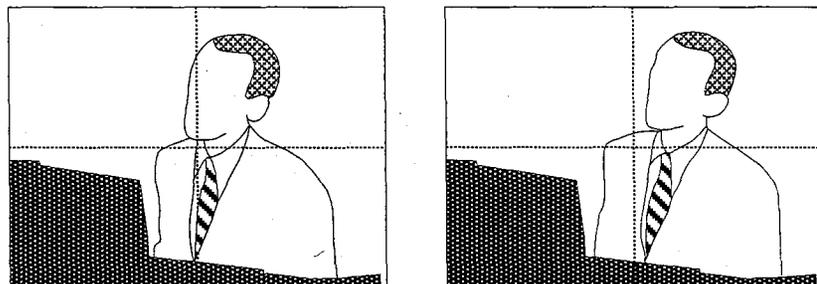
As Astaire sings 'When I'm awf'ly low' he makes a slight circular motion with his shoulders, moving first to the left of the frame (a displacement not shown in Figure 4.3) and then back past the centre, holding this position for his arrival on the long note that sets 'low'. This same basic movement is repeated as he sings 'And the world is cold'. Astaire then uses a varied version of this movement for the fourth line of lyrics: as shown in the first image of Figure 4.4, he begins



"Some-"

"day"

Figure 4.2 Astaire's gestures as he sings 'Someday'



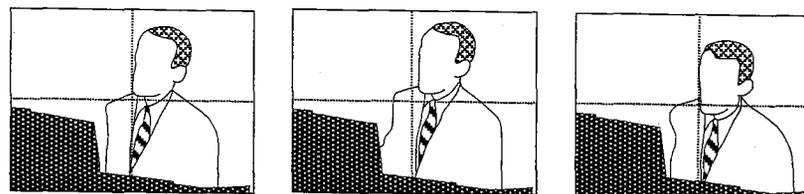
"When I'm"

"awf'ly low"

Figure 4.3 Astaire's gestures as he sings 'When I'm awf'ly low'

'I will feel' in the centre of the frame; as he sings 'a glow just thinking' he moves toward the right and top side of the frame, reaching and holding the extent of this motion on the note that sustains 'of' (shown in the second image of Figure 4.4). For the completion of the line Astaire then moves forward and dips his head, such that as he sings the word 'you' his chin reaches its lowest point in this sequence of gestures. It is worth noting that this is an expansion of the movement he made at the opening of the verse when he sang 'Someday', something that can be seen by comparing the images of Figure 4.2 with the first and last image of Figure 4.4.

For the last line of lyrics Astaire makes use of a new sort of gesture as well as returning to one he used previously. The beginning of the line ('And the') finds him dipping his head just a bit further down than it had been when he sang 'you' and moving toward the left of the frame (compare the last image of Figure 4.4 and the first of Figure 4.5). As he continues with 'way you look to-' he moves

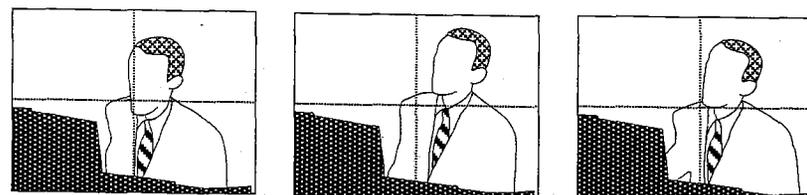


"I will feel ..."

"... of"

"you"

Figure 4.4 Astaire's gestures as he sings 'I will feel [a glow just thinking] of you'



"And the"

"way you look to-"

"night"

Figure 4.5 Astaire's gestures as he sings 'And the way you look tonight'

straight back away from the keyboard, reaching and holding the furthest point of the entire sequence. The gesture and the verse conclude as he moves back toward the keyboard, lowering both head and torso, for 'night'.

Astaire's movements as he sings the first verse of 'The Way You Look Tonight' can be analysed in terms of three gestural catchments, the occurrence and features of which I have summarized in Table 4.2. The first catchment, whose identifying feature is a dip of the head, correlates with the music of Category X. The second, whose identifying feature is a circular movement to the left of the frame and then back to centre, correlates with the music for Category Y. The third, whose identifying feature is a move away from and then back towards the keyboard, correlates with the cadential figure. Although the majority of these catchments occur singly, note that the last catchment is extended by a brief recall of the first.

Astaire's movements, then, provide a set of visuospatial images coordinated with both the music and the words of the song. They are different from the gestures studied by McNeill, however, in one important way: Astaire's movements are not part of a process through which thought is shaped, but instead *reflect* the shaping of thought accomplished through musical materials. The manifold connections between music and gesture manifested in these relationships not only speak to their connection in cognitive structure, but also point to linkages between the way a communicative medium is organized and the meaning it comes to have.

Table 4.2 An analysis of gestural catchments in Astaire's performance of 'The Way You Look Tonight'

Line	Lyrics	Gesture feature	Catchment
1	Someday,	Dip of head	C1
2	When I'm awfully low	Circular movement	C2
3	And the world is cold	Circular movement	C2
4	I will feel a glow just thinking of	Circular movement	C2
5	You	Dip of head	C1
6	And the way you look tonight	Movement away from and back towards keyboard, dip of head	C3, C1

Researchers on physical gesture generally agree that the spontaneous gesture that accompanies speech does not adhere to any sort of grammar (McNeill 2005: Chapter 2). So why would the sonic analogues of music have a grammar while the gestures that accompany speech do not? One reason is that the sonic analogues of music operate within contextual frames provided by musical rhythm and musical pitch. Rhythm – especially *metered* rhythm – offers a temporal anchor for the dynamic processes of music. For instance, each of the musical categories basic to 'The Way You Look Tonight' begins on a strong beat on an odd-numbered measure and each, save the cadential figure, ends on a strong beat on an even-numbered measure. Musical pitches – and especially *relationships* among musical pitches – provide a framework for gauging the similarity of and difference between musical events. Thus the importance of D for the melodic structure of 'The Way You Look Tonight', a pitch that earns its place as a tonal focus by its placement at key points within the tune.

The world of gesture does not typically have anything like these sorts of contextual anchors, although on occasion they do crop up. A striking example is provided by Astaire's movements as he sings 'The Way You Look Tonight'. Note the position of Astaire's chin in the second image of Figure 4.2 ('day'), the last image of Figure 4.4 ('you'), and the last image of Figure 4.5 ('night'). In Figure 4.2 and Figure 4.5 his chin ends up in almost exactly the same place; in Figure 4.4 his chin is somewhat lower, but is in approximately the same place with respect to the midline of the frame. Of course, each of these images captures an arrival point – that is, the end of a gesture. In general, there is no requirement for gestures to terminate at close to the same point in space, but this is just what happens with all three of these gestures. In the case of the tonal 'space' set up by the melody, however, these gestures conclude in *exactly* the same place: the melodic pitch D. What gesture does only rarely and approximately, then, music does frequently and with a high degree of precision, a capacity that provides a reliable contextual framework for the sonic analogues of musical grammar.

Conclusion

As I suggested in my opening comments, in many cases our characterizations of musical gesture have a metaphorical basis. On the analysis I offered there, such characterizations involve mapping structure between two domains – that is, from the source domain *physical gestures* onto the target domain *sequences of musical materials*. Astaire's gestures as he sings 'The Way You Look Tonight' suggest that this mapping can also be reversed: the sequence of musical materials set out by Kern shape the pattern and structure of Astaire's movements. From a broader perspective, my analysis of the relationship between the music and lyrics of the tune showed how musical materials, in a rather detailed way, can shape our understanding of words. The basic materials of Kern's tune – which I described in terms of cognitive categories, and which provide sonic analogues for dynamic processes – can be seen as analogous to gestural catchments, and to participate in the sort of process McNeill calls a growth point.

As research on the spontaneous gestures that accompany language has shown, gesture gives access to a dynamic, imagistic mode of thought that is inaccessible to language. In this chapter, I have endeavoured to argue that music involves a similar mode of thought, and suggested that both music and gesture draw on a common pool of cognitive resources to create analogical representations of dynamic processes. Where the analogical representations of gesture have not typically been organized into a grammar, those of music have, in part through exploiting the contextual frames provided by rhythm and pitch.

I want to emphasize that the descriptions of the relationship between music and gesture that I have offered and the outline of musical grammar that I have sketched are at a very basic level. There is much work to be done before we can understand how the dynamic processes of music give rise to the subtle and remarkable range of expression we see in the musical compositions that continue to hold our fascination. I am convinced, however, that this understanding can be based on an account of the cognitive resources humans employ in making music, resources that are closely bound to those that are also used for the gestures that accompany our speech.

References

- Duncan, S. D., Cassell, J. & Levy, E., eds. (2007). *Gesture and the Dynamic Dimension of Language: Essays in Honor of David McNeill*. Gesture Studies, Vol. 1. Philadelphia: John Benjamins Publishing Company.
- Goldberg, A. E. (2003). Constructions: A New Theoretical Approach to Language. *Trends in Cognitive Science* 7/5: 219–24.
- Goldin-Meadow, S. (2003a). *Hearing Gesture: How our Hands Help Us to Think*. Cambridge, MA: Harvard University Press.

- (2003b). *The Resilience of Language: What Gesture Creation in Deaf Children can tell us about how all Children Learn Language*. Essays in Developmental Psychology. New York: Psychology Press.
- Grady, J. E. (2007). Metaphor. In D. Geeraerts & H. Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 188–213). Oxford: Oxford University Press.
- Iverson, J. M., & Goldin-Meadow, S. (1997). What's Communication Got to Do With It? Gesture in Children Blind from Birth. *Developmental Psychology* 33/3: 453–67.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Koch, H. C. (1983[1787]). *Introductory Essay on Composition: The Mechanical Rules of Melody, Sections 3 and 4* (from *Versuch einer Anleitung zur Composition, 1782–1793*) (trans. N. K. Baker). Music Theory Translation Series. New Haven: Yale University Press.
- Lakoff, G. (1993). The Contemporary Theory of Metaphor. In A. Ortony (ed.), *Metaphor and Thought* (pp. 202–51). Cambridge: Cambridge University Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Zbikowski, L. M. (1999). Musical Coherence, Motive, and Categorization. *Music Perception* 17/1: 5–42.
- (2002). *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. New York: Oxford University Press.
- (2008a). Dance Topoi, Sonic Analogues, and Musical Grammar: Communicating with Music in the Eighteenth Century. In V. K. Agawu & D. Mirka (eds.), *Communication in Eighteenth Century Music* (pp. 283–309). New York: Cambridge University Press.
- (2008b). Metaphor and Music. In R. Gibbs, Jr (ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 502–24). Cambridge: Cambridge University Press.

Chapter 5

Distraction in Polyphonic Gesture

Anthony Gritten

Once set out a word takes wing beyond recall.
(Horace, *Epistles*)

If music drifts relatively mindless of the pressure put on it by interpretative intervention, then how might listeners engage with it? What type of aural intentionality might be appropriate if it is true that in the contemporary era 'Music escapes from musicians' (Attali 1985: 115)? How might listeners respond to the cavalier attitude of music to its own future?

Concerned with such issues around perception, intentionality and responsibility, this chapter complicates an attractive and commonsensical thesis that has been circulating for as long as people have talked about music. In the last quarter century or so, the thesis has afforded a welcome and emphatic paradigm shift in the humanities (back) towards humanity and its citizens, and in the case of music towards the ecologically grounded concerns that motivate the judgements that real listeners make and debate as part of their engagement with music. This thesis has been phrased in various ways, of which the following statement is representative: 'people can and do enjoy music without being able to make what are, in terms of musicological representation, the most elementary and basic perceptual judgements' (Cook 1990: 218; cf. Cook 1987, 1994). Acknowledging this statement's kernel of truth, I note, as Nicholas Cook did and still does, that the situation is somewhat more complex.

This chapter unpacks one issue that makes it so, working in the wake of the dialogical turn in the humanities that readings of Mikhail Bakhtin have inspired, and in the shadow of Walter Benjamin's lifelong concern to understand the 'decay of experience' that has characterized Modernity. My intention is to suggest one way of accounting for the invigorating experience of listening to musical gestures and to begin to loosen the stranglehold of an uncritical and partial appropriation of Bakhtin that has worked its way into many disciplines, musicology included. What might it mean to suggest that 'dialogic' and 'distraction' are words we live by, and what might the consequences be for understanding musical gestures?