11

Musical shape and feeling

Daniel Leech-Wilkinson

Prior (2010) and, in this volume, Prior (Chapter 7), Greasley and Prior (Chapter 8) and the many Reflections offered by practitioners all show how widely, easily and variously the concept of shape is used by musicians. And yet what do sound and shape have to do with each other? I argue in this chapter that, more than everything else it affords, shape functions as a quasi-submodal concept, common to all the senses, that readily links musical sounds to the feeling responses of listeners. Music invokes feeling states through modelling their dynamic properties, and in turn those dynamic properties are used by musicians to help them give lifelike qualities to music. In speaking of shape, musicians are indicating, by a highly efficient means, the character of, or the need for, dynamic patterns in sound that can model states of movement and feeling, or indeed of anything that changes over time. To make the case, I draw on recent work in philosophy, psychology and neuroscience in search of mechanisms capable of affording the apparent interconnectedness of shape and musical sound. The chapter builds through a survey of relevant recent work, climaxes in the middle with a view of the late work of Daniel Stern, and winds down through thought about underlying mechanisms.

Automatic and learned responses

Synaesthetes who see colours on hearing sounds sometimes report that the colours appear in specific shapes, including mobile shapes (Ward, Chapter 9 in this volume). Nonsynaesthetes, presented with animations representing synaesthetes' experiences in original and altered forms, tend to prefer the originals, suggesting that the mappings, while highly personal, are not arbitrary (Ward et al. 2008) and recruit some of the same mechanisms as normal perception (Ward, Huckstep and Tsakanikos 2006). That a music–shape relationship is used so widely by musicians (Prior 2010, 2012) suggests that the mapping is very easy to make, and therefore that the capacity is to some extent structural or at any rate very thoroughly learned.

Walker et al. (2010) showed that infants as young as three or four months prefer matchings of visual and pitch direction (both moving up, or both moving down) and pitch and sharpness (rising and pointed, or falling and rounded) that correspond to the mappings reported in the extensive literature on adult cross-modal mapping and synaesthetes, implying that infant perception might be synaesthetic at birth though, in most people, later unlearned (Maurer and Mondloch, 2006). But how fully unlearned?

Ramachandran and Hubbard famously demonstrated (2001; see also Maurer, Pathman and Mondloch 2006) that almost everyone agrees that rounded and pointed shapes map onto the nonsense words 'bouba' and 'kiki' rather than 'kiki' and 'bouba', confirming the findings of a similar experiment by Köhler (1929) with 'baluma' and 'takete' (see Spence 2011 for a review of these and other studies, including cross-cultural ones). Exceptions have been found among children with autism spectrum disorder and people with damage to certain brain areas, suggesting 'that crossmodal correspondences, at least those involving sound symbolism, can occur at quite a high level' (ibid.: 974).

On the other hand, Dolscheid et al. (2013, discussed from a similar perspective in Eitan 2013) provided good evidence that the metaphors for pitch position available in one's native language influence one's sense of pitch. For Dutch speakers, high and low visual stimuli influenced the pitches they sang, while thin and thick did not, and for Farsi speakers, for whom pitches are thin and thick rather than high and low, it was the thickness of the visual stimuli that had the effect, to approximately the same degree. In this respect, at least, language use apparently feeds back into music cognition. Training Dutch speakers in the use of the Farsi terms produced a similar result, but training them in the opposite terms (thick for high and thin for low) did not. It seems, then, that language has an effect only in the direction already established in pre-linguistic infant synaesthesia: it can reinforce but not fully determine our responses (Eitan 2013). Similarly, the linguistic terms (high, low, thick, thin) appear to have been adopted by languages because they have pre-linguistic origins.

Many of the synonyms for musical shape offered by participants in the research leading to Prior (2012) can be classified linguistically (though not necessarily according to participants' intentions) as drawing on images of either intensity/quantity or trajectory/direction or both (Table 11.1). Images suggesting simple mapping onto two-dimensional space—shapes as visualized forms—are less frequent than those suggesting motion with change, which is perhaps not surprising given the nature of music, but it nonetheless suggests that there is much more going on here than simply the pitch content of music being imagined as height in relation to time. (More sophisticated uses are examined by Leech-Wilkinson and Prior 2014.)

TABLE 11.1 **Some of the synonyms for 'shape' collected for Prior (2010)**

Quantity	Both	Direction
Large and small	Expansion and contraction	High and low
Light and shade	Growth and decay	Backwards and forwards
Fast and slow	Swell and dying away	In and out
Thick and thin	Ebb and flow	Up and down
Give and take	Peaks and troughs	Left to right
Anticipation	Crest and trough	Ascend and descend
Tension	Push and ease	Driving and following
Intensity	Stretching and relaxing	Movement
Climax	Build and release	Flow
Contrast	Impetus	Moulding
Timing or pacing	Urge	Direction or energy
Change over time	Impact	Momentum
Rhythm	Dynamics	Growth

Many studies have examined pitch/height perception, and taken together they offer considerable support for the idea that this is a relationship that is stronger for those used to seeing music notated with pitch on the y axis and time on the x than for others (Eitan and Granot 2006; Antovic 2009, comparing Serb and Romani children; Athanasopoulos and Moran 2013, comparing British, Japanese and BenaBena; Küssner and Leech-Wilkinson 2014; Küssner et al. 2014). At the same time, the notion that pitch has height may, as we have seen in reporting Walker et al. (2010), rest on deeper foundations. Prince, Schmuckler and Thompson (2009: 42), surveying previous studies, report that 'both musicians and untrained listeners exhibit activation in visual cortex while attending to pitches within a melody', but they go on to find that 'converting contour information from the auditory to the visual domain exploits the skills that musical training confers'. In other words, musical training sharpens a sense that pitch maps onto space, but (just as we saw in assessing the role of language) this particular cross-modal mapping may call on a more widespread ability.

Perhaps a more challenging question to ask is how the dimensions of music other than pitch are imagined in terms of space. Can we extend the notion of shape in its visuo-spatial sense to more aspects of musical sound? And if so, can they be combined to generate the strength of association that we seem to find 'shape' having for many musicians?

Eitan and Granot (2006) investigated perceived similarities between several musical parameters (including 'dynamics, pitch contour, pitch intervals, attack rate and articulation') and found human motion associated with them all, though in varying degrees and directions, as summarized in Table 11.2. The finding on speed is revealing. Acceleration might associate with descent

TABLE 11.2 **Associations reported in Eitan and Granot (2006)**

Loudness	
Crescendo	Coming closer, acceleration, increasing energy (due to external force in slow tempos only), running motion
Diminuendo	Moving away, deceleration (at slower tempos only), falling pitch, falling or sliding motion
Pitch contour	
Rise	Acceleration, spatial ascent, moving away (small effect), higher energy, running or walking
Fall	Deceleration, spatial descent, lower energy, leftwards motion, falling
Speed	
Acceleration	Descent
Deceleration	Descent, moving away
Articulation	
Legato changing to staccato	Moving away, slowing (at slow speeds)

Note: Findings listed here combine some of their categories.

through our experience of gravity, whereas deceleration might associate with descent through the quite different experience of increasing exhaustion or declining intensity of feeling or engagement. The embodiment of music is complex and multifaceted, as we can see from the subtlety and accuracy with which music can associate with other phenomena (Leech-Wilkinson 2009a; Juslin and Lindström 2010).

While Eitan and Granot (2006) found that associations involving verticality and speed change are stronger for musicians (i.e. those with musical training) than for nonmusicians (those without), other associations—including loudness with motion and energy, and pitch direction with energy—seem equally strong for both groups. Küssner and Leech-Wilkinson (2014) found 83 per cent of nonmusicians matching pitch and height as against 98 per cent of musicians, and 73 per cent of nonmusicians matching loudness and energy as against 93 per cent of musicians. The 73 per cent figure for the nonmusicians is not trivial, even though their responses are more varied. It is hard to say, then, that shape is linked to pitch height (still less loudness to energy) largely through training. However strengthened the association may have become through practice, it seems distinctly possible that it has deeper roots, though whether biological (inherited) or embodied (acquired through physiological interaction with the environment, Johnson's sense of 'embodied'; 2007) remains an open question. A sense of musical shape, then, may have more to do with the multiply changing dynamics of an event than with its pitch contour (though that too is important for musicians). The importance of this will become clear when we look at Stern's work on the dynamics of everyday life.

So far we have considered relationships between visualizable shape and music. But what of other shaped experiences? Intensity, which is found in all dimensions of musical sound, was identified by Marks (1978) and Eitan and Granot (2006) as a key attribute through which we associate auditory and nonauditory domains, while Eitan and Granot (2007) showed how easily it is mapped across musical parameters. Stevens et al. (2009: 806, 809) suggest that arousal response is more consistent than valence, which may be more subject to cultural norms, which in turn suggests that response to intensity (in so far as arousal is intensity) may be the more fundamental mechanism. We may go further and suggest that intensity comes closest (among qualities accessible to conscious awareness) to the unidentified common parameter among the terms Eitan and Timmers (2010) collected for pitch differentiation: the young, the high, the sharp, the bright, the fast, the strong, the awake, the active provide us with more intense experiences (i.e. greater disruption of a comfortable state) than the old, the low, the blunt, the dark, the slow, the weak, the sleepy, the passive. Intensity maps easily onto these as well as onto musical components, and also onto the underlying processes of tension, surprise, disruption and change that are fundamental to that sense of directed motion found in musical structures and performances and, indeed, everyday life.

Most of the studies cited so far, and many more, give us good reason to think of music as lifelike. That is to say, how sounds interrelate (with what was just heard, with what happened earlier, with what happens next in the sense that in the light of that they seem more meaningful) and the dynamics of the succession of sounds cause them to seem alive and construct a vivid sense of 'now' that seems to move with them through time (Johnson 2007: MOVING TIMES schema; Stern 2004). Music is lifelike, then, through the constantly changing shapes of the experiences it engenders in a listener. It is lifelike, too, through our tendency to chunk events into narratives, separate from the events around them (Lavy 2001; Murray 2003). It is not that music does anything as crude as to tell a story, but rather that as one chunk or phrase or event is followed by or transformed into the next, a sense of narrative develops which in its own musical terms is coherent and directed (by the composer or performer or both, in collaboration with the listener). The narrative has shape at a higher structural level than the shapeliness of moment-to-moment progressions; but at the moment-to-moment level it is made up of lifelike behaviour, its life played out in sounds and in our responses to them.

Both music and life can be summarized formally as narrative structures, but both are experienced in time through sequences of events whose most striking characteristic is their vitality, created through the manner in which they change (Stern 2004). Experiences (the events that in sequence accumulate into the narrative of our lives) begin either gradually or suddenly, and typically develop either by building to a point of maximum engagement and then tailing off, or just by tailing off gradually as one deals with the consequences of a sudden start. In this important sense experiences have shape; indeed, that they are shaped is one of their most evident characteristics. What shapes them are changes in all aspects of perception and response as they develop. Because it is constantly changing its dynamic characteristics, and because of the ease with which it maps across domains, music models or at any rate behaves like the experience of events. And we map ourselves onto it, experiencing it like a sequence of lived events with the feelings they arouse (Johnson 2007: esp. 236). In a number of respects, then, music shapes and is shaped through the dynamics formed by its changing sound.

If we look at music from these perspectives, it becomes very easy to sense why music and shape are so closely associated, but it is also clear that this is less directly a matter of shapes that may be visualized in space (though one could make drawings or diagrams of many of these things) and more a matter of changing shapes of feeling, that is to say, change in the intensity and quality of experience over time. Susanne Langer seems to have been speaking of a relationship between shapes of feeling and shapes of music when she famously concluded that music's intrinsic properties '*sound* the way moods *feel*' (Langer 1963: 244–5; her italics). And so to make progress in understanding the ease with which the dynamic content of music maps to and from that of lived experience, in other words for an effective approach to analysing musical shape, we need a theory that deals with the psychodynamics of everyday life. We find one in the work of Daniel Stern. (On Stern, see also Kim 2012, 2013.)

The dynamics of experience

Stern's final work is concerned with *The Present Moment* (Stern 2004) and with dynamic experience or *Forms of Vitality* (Stern 2010). Phenomenology has long had related concerns, but Stern's particular focus is on the experience of 'now', the ever-present moment. Stern sees experience as characterized by a sequence of present moments, each no more than a few seconds in length, which are shaped by feeling responses to incoming perceptions, and which group together to form dynamically shaped mini-dramas, sensed as a gestalt, through which one lives. Stern records examples for analysis by asking participants to describe episodes in their everyday lives. Here, a participant describes breakfast.

> *Present Moment 4*
> I am holding a slice of bread, not yet spread with honey. But it is a different kind of bread than I normally buy. It feels strange and I am surprised by it. I think, 'What do I do with this bread?' A mild negative feeling arises.

This moment took about three seconds. She then spreads honey on the bread without paying conscious attention to the act. A new moment begins adjacent to the previous one.

> *Present Moment 5*
> I am then aware of biting into the honeyed bread. I like the texture and think, 'It's not so bad.' And with that a sense of feeling better builds up. I then become conscious of the radio interview again. (Stern 2004: 13–14)

What we have here is a sequence of remembered experiences split into their constituent moments; each little scene has a thematic unity, dealing with one main event plus the thoughts and background events that accompanied it. The description is necessarily constructed after the event—to self-report simultaneously would be to alter the experience even more—and therefore an element of composition is inevitably involved. In a sense, this is not unhelpful for a comparison with music, where, in the absence of semantic meaning, narratives are often consciously constructed as part of the process of finding meaning in sounds (Wingstedt, Brändström and Berg 2010; Tarasti 2004). Stern provides a map of the feeling shapes that constitute these two 'moments', using curves to represent the changing intensity of each feeling shape ('surprise', 'feel negative', 'bite + chew', 'feel resistance', 'surprise', 'pleasure'), using the changing shape and thickness of the curve to give an impressionistic sense of what the participant reported having experienced (Stern 2004: 15).

Stern derives from his case studies conclusions about the present moment that are very easy to apply to our experience of music.

- Present moments [read 'musical phrases'] are unbelievably rich. Much happens, even though they last only a short time.
- Present moments occupy the subjective now. . .
- The moment is a whole happening, a gestalt. The psychological subject matter is the whole, not the smaller units that make it up.
- When she experienced these events it occurred in a now that she could identify and put boundaries around.
- The present moment is short. In this case . . . each lasted between three and five seconds, as estimated by the subject.
- Consciousness is the main criterion used to identify episodes containing present moments. . .
- The feelings experienced . . . trace a time-shape (a temporal profile) of analogic risings and fallings. In other words, they are carried on vitality affects (dynamic time-shapes) that contour the experience temporally.
- A lived story unfolds within each present moment. It is made of many small experiences that are put together in the subjective present. The storyline, even if minimal, rides on the temporal feeling shape of the

contoured affects. The unfolding micro-story resolves the novelty or problem.

- Such moments are not cut off from the rest of life, isolated and unconnected. Rather, they capture a sense of the subject's style, personality, preoccupations, or conflict—in other words, their experience of the past. Each such moment is psychodynamically relevant. (2004: 14–16)

So too music, when one focuses one's attention on it, can be experienced as wholly engrossing, 'unbelievably rich'. It occupies the subjective now. It creates a sense of stylistic and psychological wholeness (the one leading to the other). Its phrases and sections have definition and gestalt-like qualities of self-containedness within a greater whole; and at the local level they can be quite short, short enough to be experienced within a psychological 'now'. We identify and understand music in terms of these experienced 'moments' which give it its meaning for us. The feelings it generates 'trace a time-shape of analogic risings and fallings. In other words, they are carried on vitality affects (dynamic time-shapes) that contour the experience temporally'. We sense musical continuity as having narrative qualities, although it usefully lacks the precision of the 'lived stories' which Stern finds make up everyday life. They 'capture a sense of the [music]'s style' and analogously behave as if they were people, indeed as if they were us. Thus musical units are indeed 'psychodynamically relevant'.

Similarly, the curves Stern uses to represent these feeling shapes could equally well be mapping aspects of a phrase, or melody, or loudnesses, texture, rhythm, or also qualities like expectedness, complexity, mood, character, edginess, tension, all adding up in complex ways to give a sense of shape. This itself emphasizes how directly the musical features (contour, loudness, speed) model the qualitative (mood, tension and so on).

Stern's 'present moments' bring together things one experiences as 'now' in time spans that are only as long as the 2–5 seconds our perceptual systems allow (Fraisse 1984). Stern here draws on Husserl's phenomenology of time: the past of the present moment, not remembered but rather still experienced because still fading from the present; the present of the present moment; and the anticipated, expected, immediate future. This ties in usefully with existing music perception and analytical theory, which argues that expectation is essential in the creation of musical meaning (Huron 2006; Stern's particular reference point is Narmour 1990). But Stern's point is that all three stages are taken together as 'now'. This offers a more ecologically plausible notion of what it is like to hear and to perform music, both performer and listener aiming always to treat what comes next as a good continuation of what just happened. It follows, both for Stern describing life and for us thinking of music, that the present moment is not always the most intense; it may be happening in the shadow of what happened a moment ago, but what happened a moment ago is also something that can define 'now' and give it shape. Indeed, 'now' can be relatively trivial so long as one understands it in relation to the past and the anticipated future. And Stern goes on to make precisely this analogy with music, noting how 'much of the richness of music lies in the fact that each subsequent phrase recontextualizes the previous one', but at the same time, 'A coherent experience was grasped during the present moment, even though that experience may have multiple fates' (2004: 30). There is thus a constant 'trialogue' between past, present and future characteristic of music and life (ibid.: 31).

But there are important differences. Music has hierarchical levels, for one thing. It has an orderliness, in other words, that we do not find in life. In music each event leads coherently into the next, or contrasts with it in a way that will later be reconciled or resolved. There are no loose ends in music, or if there are we tend to fault it, saying it is badly composed. Music is lifelike in a utopian fantasy world where everything that happens makes sense and where, whatever conflicts we may experience along the way, everything turns out for the best. The entire tradition of music analysis is directed at proving this. (On music as utopia, see also Levitas 2010.)

In Stern's view, present moments characterized by 'temporally contoured feelings' (2004: 36) are best thought of as 'vitality affects': 'these temporal contours of stimulations . . . are transposed into contours of feelings in us' (ibid.: 64). Thus 'temporal contours' are the objective changes in intensity or quality of the stimulations; 'vitality affects' are the subjectively experienced shifts in internal feeling states that accompany the temporal contour, their 'vitality' being that sense in which they are the most characteristic aspect of the affective experience of living. Most stimulation of the nervous system, Stern says, whether it comes from within or without, 'has a temporal shape or contour that consists of analogic shifts in the intensity, rhythm, or form of the stimulus' (ibid.: 62). And this, of course, is exactly what happens in the performance of a (western classical) score. Through the score, a roadmap is already provided by the composer—the events of a life are determined—but what the performer does is precisely to provide a temporal shape through moment-to-moment adjustments in intensity, rhythm and pitch (or whichever of these dimensions is available through the instrument in use) that provide triggers for the feelings that seem best (most satisfyingly) to arise from those events. One can think of these as 'expressive gestures' (Leech-Wilkinson 2006, 2009a). Expressive gestures generate vitality affects: it is through expressive gestures that musical scores, which without them are so dull as barely to be music at all,[1] come to life. And what musicians overwhelmingly mean when they talk about 'shaping' a phrase given in a score is, as we have seen, that process of enlivening that is achieved through the use of expressive gestures.

For Stern, though, the key idea is that the dynamic qualities of vitality affects may be linked across modalities (Stern 2004: 37, 64–5). Vitality affects occur in many modes and can map easily from one mode to another. This is certainly what has been repeatedly suggested for expressive gestures, which index other things, or behave like them, and take meaning from the likeness (Leech-Wilkinson 2006). As Kim (2013) points out (in her essay for the special issues of *Empirical Musicology Review* on music and shape), for both Stern and his predecessors, Hausegger (1887) and Truslit (1938), the experience of expressive shape (vitality affect) is 'understood in a broader sense than emotions' (Kim 2013: 165). The emphasis here is on dynamics, not states, and in that sense (shaped) forms of vitality may offer a more appropriate way of thinking about the experience of music than seeing music as a sequence of inductions or representations of emotional states (ibid.). A shape-focused view of music, therefore, might offer a more ecologically valid way of understanding feeling responses to music than do attempts to see music as expressive of particular emotional states.

A final point of Stern's (2004), though he intended it as a key to psychoanalytic practice, ties in with Kim's (2013) focus on the aesthetics of empathy and is especially relevant to our understanding of musical performance and response. For Stern (2004: 22, 172–3) present moments shared, where analyst and analysand seem to understand each other's feelings with unmediated clarity, can be especially intense and life-changing. Stern describes these as 'a shared feeling voyage', when 'two people traverse together a feeling-landscape as it unfolds in real time.' It seems possible that something like this takes place between musicians performing well together and is imagined as happening also for listeners attending to a performer with exceptional concentration and sympathy. In the latter case, the listener, and perhaps the performer, feels that they understand the other and become one with their music.[2] It may be an illusion, but there is intersubjectivity concentrated in these moments, as music happens, of a sort that may not be unrelated (though this needs focused research) to song as sexual attractor (Miller 2000).

Stern (2010) takes the notion of vitality form (a more abstract conception of the form underlying a vitality affect) and sees it now as the most fundamental percept characteristic of life: 'Subjectively, a thought can rush onto the mental stage and swell, or it can quietly just appear and then fade. It has a beginning, middle, and ending. . . Mental movement, while it is happening, traces a profile of its rising and falling strength as it is contoured in time. This is its dynamic form of vitality' (ibid.: 21). At the same time,

> Vitality forms are hard to grasp because we experience them in almost all waking activities. They are obscured by the felt quality of emotions as it accompanies them. They are absorbed into the explicit meaning as the vitality form accompanies a train of thought, so we do not pay attention to the feel of the emergence of the thought, but only to its contents. It [the vitality form] slips through our fingers. (ibid.: 10)

And so it is with musical shape. Our attention is so focused on what is happening among the notes—harmonies, rhythms, themes, events, qualitative effects—that we normally pay no attention to the shapes traced by these sounds and by the feelings they engender. Yet it is their shape that carries all these experiences we notice so directly. This perhaps begins to explain why musical shape, until recently, lacked a bibliography. On the one hand, as an attribute of the attention-grabbing musical surface, it is rarely noticed for itself; on the other, among musicians the term is so easily used as a shorthand for expressivity that it hardly needs explanation for it to work perfectly well as a heuristic (Leech-Wilkinson and Prior 2014).

As far as the underlying mechanism is concerned, 'Once an experience *activates* the brain, it will leave a purely vitality dynamic representation and a content representation.' (The content representation refers to the causes of the stimulus and their implications for us.) 'The dynamic representation must encode the speed and its changes, the intensity (force) and its changes, and the duration, and the temporal stresses, rhythm, and directionality' (Stern 2010: 25). So '[v]itality forms are modality non-specific. They belong to no one sensory modality but to all. . .' (ibid.: 26).

This helps to explain why music expresses vitality forms so easily and so clearly. Like vitality forms, music's modality is also in important senses non-specific; it can apply to anything that changes over time or (if we wish) to little beyond itself. And within this abstract medium of music the notion of shape is abstract too: it offers a concept common to, and realizable through, many aspects of music, just as music can model many aspects of life. In this sense, shape is close to being a suprasensory attribute, linked to the most evident of all suprasensory modalities, which (Marks 1978 argued) is intensity. But where intensity characterizes a moment, shape summarizes a process, and it is that that gives it outstanding heuristic value for musicians. So we can say that shape describes the changing intensities that model one kind of experience in the domain of another. For Stern, 'The concept of dynamic vitality forms brings together four converging lines of thought, namely intersubjectivity, cross- and meta-modality, the dynamic features of experience, and a phenomenological focus on subjectivity' (2010: 44). And these are exactly our concerns too. Musical shape communicates among performers and listeners; it maps between modalities and can be conceived as beyond them; as a concept it encapsulates a fundamental aspect of lived experience; and it characterizes everything we feel happening within ourselves.

Stern's work, then, helps to explain how it is that the notion of shaping music belongs not just to musical practice, nor to a particular tradition in which, thanks to notation, pitch is conceived as moving up and down through space. Rather, it offers a way of understanding how it is that music maps so easily onto our experience of changing affect. Like musical sound, our feelings change in intensity, in complex polyphonic patterns, from moment to moment, experienced always as 'now' with a brief, still-vivid past (preceded by a longer, increasingly hazy tail) and an implied immediate future which may yet surprise us. Shape conceptualizes all these aspects of music and feeling. More than that, it carries them along in us. As Stern emphasizes, his work has significant commonalities with phenomenology, sharing its prioritizing of experience over structure or language as the dominant reference domain for cross-modal signification. What seems to be implied by the prioritizing of lived experience is that knowledge of music acquired through our bodies and their preconscious, unreflective motor and limbic responses underlies cognitive responses, providing a frame within which they operate and in relation to which they are selected and form as thoughts about music. Underlying Stern's work, therefore, and also the work on embodiment that we consider next, is the wealth of recent research on music's interaction with the brain's limbic and motor systems (Koelsch 2010; Altenmüller, Wiesendanger and Kesselring 2006). A full understanding of music and shape, in so far as present research allows, would need to take detailed account of that work.

Embodiment

While the foundation provided by underlying neural mechanisms is acknowledged, to understand shape in terms of musical experience we need an approach to thinking about music and bodily response that illuminates particular cross-domain mappings. Stern offers one such approach. Another valuable source for thinking about musical shape is the ever-growing literature on embodiment. Mark Johnson's *The Meaning of the Body* (2007) is particularly helpful here. Johnson argues that the movement of our bodies through our environment, and our daily experience of other moving objects, are fundamental to our understanding of the world and provide us with ways of thinking about everything that involves a sense of motion: 'For example, *tension* has a meaning grounded in bodily exertion and felt muscular tension. *Linearity* derives its meaning from the spatial, directional qualities of bodily motion. *Amplitude* is meaningful to us first and foremost as a bodily phenomenon of expansion and contraction in the range of motion' (Johnson 2007: 25).

These are not just abstract qualities: 'they are *qualities of organism–environment interactions*' (ibid.). We might add that they are essential features of music and that they are also qualities in which change can be experienced more holistically as shape. Drawing on Johnson's discussion of Dewey's theory of qualities—the overriding feeling of a situation—which Johnson sees as being created by a collection of cross-domain mappings (ibid.: 77), one may go on to propose that we understand music crucially (though not wholly) as the constantly changing, dynamic, felt quality created by cross-domain mapping from the sounds of a particular performance. This seems to make sense of everyday experience of music. The quality arising from embodiment is modulated, of course, by many additional factors including all kinds of autobiographical and learned associations and intellectual thought about the music. But it remains a crucial ingredient in the mix.

With George Lakoff, Johnson had already done important work on the schemas that we use in thinking about our interaction with the environment (Lakoff and Johnson 1980), and a number of those that Johnson sees as crucial to the construction of meaning through embodiment are closely concerned with aspects of musical shape. Especially relevant is SCALARITY: 'Because we must continually monitor our own changing bodily states, we are exquisitely attuned to changes in degree, intensity and quality of feelings. Such experiences are the basis for our sense of the scalar intensity of a quality (the SCALARITY schema)' (Johnson 2007: 138). What we gain by relating shape to the SCALARITY schema is an increased sense of the multiplicity of ways in which changing intensity of feeling, arising from changing intensity of musical score and performance, acquires further depth and affords varied meanings from the ease with which a performance maps onto many other aspects of experience. It is not necessary to assume that we inherit these schemas: we simply cannot fail, given the bodies we have and the environment in which we find ourselves, to learn them (ibid.: 178). Embodied schemas may well form at a level between the biological and the cultural: embodied knowledge is neither inherited nor indoctrinated; it is acquired simply by living and moving.

This helps us tease out the difference between schemas, like SCALARITY or MORE IS UP / LESS IS DOWN, which are embodied, and the association that is probably learned by western musicians between pitch and height, which for them is also connected with shape. The sight of notated sounds rising and falling across the page, repeated daily from childhood, will inevitably create a deep-seated sense of rising and falling pitch which would give the notion of shape one particular intensifying aspect for western classical musicians that it might not have everywhere (as found by Athanasopoulos and Moran 2013). To western musicians it would come to seem as natural as a schema and would increase the sense that music is shaped by verticality as well as by intensity.

Further image schemas that bear on musical shape include MOVING MUSIC with its underlying schemas MOVING TIME (we imagine time, and music, moving past us) and MOVING OBSERVER (we imagine ourselves moving through time with the musical now moment), the latter with its associate MUSIC AS MOVING FORCE; and also MUSICAL LANDSCAPE (we imagine a performance of a score as simultaneously present, spread out around or before us, the schema underlying images of musical compositions as formal structures fully present independent of performances; Johnson 2007: 29–31, 244–54; Johnson and Larson 2003). Shape is used by musicians, as Prior (2010, 2011) has shown, through all these notions. Seen as ways of thinking about music, they can easily seem contradictory, even mutually exclusive. How can you think of music as moving force and also as landscape? Understanding the rootedness of these conceptualizations of music in low-level image schemas, each with a firm basis in embodied experience, explains easily how all can be drawn on in understanding music. Rooted in our bodies, they each arise reasonably from our daily experience, and as such there is no functional conflict between them (Johnson and Larson 2003: 80).

Underlying principles

Stern and Johnson are in many ways offering similar views, albeit with differing foci.[3] Stern is concerned with the dynamics of feelings as they are experienced phenomenologically, Johnson with the process by which those feelings have meaning. Together they offer us the beginnings of a coherent explanation for the meaningful interaction of music and shape. They also suggest how it is that shape seems to represent so many aspects of our experience of music both as sound and as meaning—melodic contour, harmonic tension, loudness envelopes, pitch inflections, textural change, performance actions, affective (in the right social context, physical) response—most of which can function simultaneously on many structural levels, shapes nested within shapes within shapes and so on from a single note to a whole piece. It may be possible to take our understanding of this phenomenon one step further, however, by going back to some of the research on which Stern and Johnson draw, and supplementing it with findings from more recent work.

Two research themes are of particular interest: suprasensory modalities and multisensory perception. Both offer ways of understanding the mechanics of cross-domain mapping. First, though, we need to consider the level on which mapping takes place. As Stern points out, Lawrence Marks in his classic study of the theories then relevant to synaesthesia, *The Unity of the Senses: Interrelations among the Modalities* (1978), traced ideas about suprasensory attributes as far back as Democritus (c. 460–370 BCE) and Aristotle (384–22 BCE). For Aristotle a sixth, common sense, integrated the outputs of all the others, allowing them to be perceived in terms of general attributes: 'motion, rest, number, form, magnitude, and unity' (Marks 1978: 4; Doğantan-Dack 2013 points out a predecessor in Ehrenfels). Marks proposed that a level on which (perhaps due to a common phylogenetic heritage) the senses overlap, and where they could find features in common, could explain the ease with which we make analogies between different sensory experiences (ibid.: 182–5). To clarify what kinds of common features these might be, consider the process of analysing auditory data into perceptual features—including melody, contour, rhythm—demonstrated by the results of selective brain injury (Peretz 2003). This necessarily involves the process, already well documented in studies of auditory perception (summarized in Rees and Palmer 2010), by which the auditory cortex analyses incoming sound. The question is what happens between that initial analysis and our perception of melody, contour, rhythm and so on, synthesized out of them.

It is easy to see that contour, for example, combines aspects of frequency, relative quantity, speed and time, and is already quite a complex phenomenon. Frequency, quantity, speed and time are not entirely straightforward either, perhaps involving signals registered in a simpler form. So in constructing a sense of contour the brain has to analyse out of the incoming sound data, and then combine into a contour gestalt, components that are very much simpler. Alternatively, if the gestalt is perceived first (which is also possible) it may then be deconstructed into its components in order to search for them in gestalts experienced before, identifying them and giving them meaning through previous experience. Either way, in their simplest, lowest-level form, we have no perceptual access to these simpler components. We perceive them only within more complex phenomena. However they are registered, each has a role in other percepts too: frequency and quantity are necessary components of loudness, for example, since without frequencies there is nothing to be heard, and without a variable response to quantity there is no way of loudness being a percept. These simpler components would not be consciously perceptible on their own, but precisely because they are so simple, not yet combined into specific percepts, they could be simple enough to be shared across several (or all) sensory modalities. All that is then required for cross-domain mapping is a mechanism in the brain that habitually compares one data stream with, or that has a response mechanism sensitive to, a memory of others.

Both the simpler components and the mapping mechanism are implied in studies by, respectively, Näätänen and Winkler (1999) and McLachlan and colleagues (2010, 2011). Näätänen and Winkler propose 'sensory feature traces' assembled during a pre-representational phase of sensory information processing which then, in the following representational phase, are mapped onto time and compared with the contents of long-term memory. It would be at this second stage that features such as frequency, relative quantity, speed and time become perceptible. First-stage features, Näätänen and Winkler suggest, are simpler and inaccessible to conscious perception, which is exactly what we need to explain cross-modal commonalities. Refining Näätänen and Winkler's work in a major survey of recent research on auditory processing, McLachlan and Wilson (2010) propose that incoming auditory information is stored in a multidimensional array in short-term memory, enabling multimodal similarities with data in long-term memory to be identified and to contribute to sound source identification and association (further developed in McLachlan et al. 2011)

This process, or something like it, allows us to make sense of the very high rates of agreement that Eitan and Timmers (2010) found among participants faced with having to choose which of two terms was most like what in the West we call 'high' pitch. Table 11.3 shows the terms that produced most agreement. I have added a column proposing an association learned through everyday

TABLE 11.3 **Highest-scoring results from Eitan and Timmers (2010; Table 1), with a proposed environmental cause for the participants' preference**

Highest-scoring Test Terms for 'High' Pitch	Consensus High	Consensus Low	Experience Underlying Preference
Alert–sleepy	1.0	.10	voice
Sharp–heavy	.98	.00	size
Thin–thick	.98	.03	size
Sharp–blunt	.97	.08	struck sound?
Young–old	.97	.10	age
Happy–sad	.95	.03	voice
Light–dark	.95	.05	?
Fast–slow	.95	.05	size
Light–heavy	.95	.11	size
Granddaughter–grandmother	.93	.08	age
Feminine–masculine	.90	.07	gender
Small–large	.90	.15	size
Tense–relaxed	.90	.17	body

experience which may underlie the participants' preferences. Thus, old people have lower voices than young, so high pitch is female or a granddaughter, while low is male or a grandmother and so on. But other examples have no such everyday explanation. What has pitch to do with brightness in our experience of the world (an association we may share with chimpanzees: Ludwig, Adachi and Matsuzawa 2011, though see also Spence and Deroy 2012)? In this case, it seems necessary to hypothesize that an aspect of pitch, too low-level for us to perceive, is being found to correspond to (or even be identical to) an aspect of light. And the analysis of, for example, sound and light into very basic components seems necessary for the construction of the gestalts evidenced by the selective results of brain injuries (Peretz 2003) and for the process suggested by Näätänen and Winkler (1999).

The same mechanism seems likely to underlie Marks' (1978) suprasensory attributes: 'Suprasensory attributes are those categories or dimensions of experience that . . . apply to most or to all modalities. Intensity is a classic example, to which duration must also be added. Size (extension), brightness, and hedonic tone are other candidates, though perhaps not universally applicable' (ibid.: 5). As we have seen in discussing automatic and learned responses, intensity comes closest to the generalized quality that seems necessary to link pitch and brightness, for example. Nonetheless, it may already be too complex to be fundamental to cross-domain mapping. Shape–sound relations seem to depend less on intensity as a basic category than on a lower-level aspect of changing quantity that can behave identically in awareness of physical space (allowing sound to be sensed as having height or proximity) and of intensity of feeling.

Bueti and Walsh (2009) offer a mechanism by which a sense of magnitude may be acquired from infancy, building on Walsh's (2003) work on quantity. But all these could be synthesized out of Näätänen and Winkler's (1999) 'sensory feature traces'.

Martino and Marks (2000) introduce key additional points. First, relationships between modalities depend on an awareness of the relative position of a stimulus on a scale defined by experience. We can say that a sound is bright only because we can compare it to other sounds we have heard that are darker. No sound can seem one or the other without that experience being accessible to us. Here is where synaesthesia differs from cross-modal experience, however, since for synaesthetes these relationships are automatic and invariant. This relativity of attributes for most listeners is entirely compatible with the view I have suggested: comparison is made between incoming data and things already known (whether known through inheritance or, much more usually, through embodiment, enculturation or learning). There is no need to assume measurement against a baseline.

Secondly, Martino and Marks point out that a post-perceptual representation of stimuli (analysis into constituent parts and cross-domain mapping *after* perception of domain-specific gestalts) would explain recent findings that semantic relationships between stimuli can also produce congruence effects (cross-domain mappings). McLachlan and Wilson (2010: 181) propose a similar role for verbal labels at an early stage in sound identification. Again, this allows (indeed requires) embodiment, enculturation or learning to play a powerful role in establishing mappings: experience creates rapid matching of stimulus and meaning involving whichever domains have repeatedly been found relevant (Yu 2008). Thus, music getting quieter maps to increasing distance because of repeated experience in the real world (embodiment); music increasing in frequency maps to increasing height because of repeated linguistic experience of 'high' and 'low' as descriptors of musical sounds (enculturation), or more specifically for western classical musicians because of repeated experience of notation (learning). Or to take a different kind of example, violent music is embodied, martial music is encultured, cigar music is learned (see www.youtube.com/watch?v=NIckHmwZAeI).[4]

The other very relevant body of work that we need to consider (as Stern 2010: 49 suggests) concerns the possible coupling of motion and thought within the sensorimotor system. Gallese and Lakoff (2005), building on Gallese's earlier work on mirror neurons and Lakoff's on embodied concepts, offer an explanation for the strength of couplings that bring together emotional response with thought and motion. Musical shape, which involves all three, would provide an outstanding example. They argue that 'sensory modalities like vision, touch, hearing and so on are actually integrated with each other *and* with motor control and planning' (ibid.: 459). The same areas that control action also construct 'an integrated representation of (1) actions together with (2) objects acted on and (3) locations toward which actions are directed' (ibid.: 460). In other words, they argue that action, imagined action and understanding are all tied together through using the same neural systems. The same might very well be true for musical performance, listening to musical performance and understanding it. A very similar point is argued by Molnar-Szakacs and Overy (2006: 236): 'according to the simulation mechanism implemented by the human mirror neuron system, a similar or equivalent motor network is engaged by someone listening to singing/drumming as the motor network engaged by the actual singer/drummer; from the large-scale movements of different notes to the tiny, subtle movements of different timbres.'

Recent work on multisensory perception offers further ways in which sound might generate a sense of shape. A review by Stein and Stanford (2008) concludes that many or even most neural systems may be multimodal. One of their sources, Ghazanfar and Schroeder (2006: 284), seems especially pertinent:

> Traditionally, it has been assumed that the integration of such disparate information at the cortical level was the task of specialized, higher-order association areas of the neocortex. In stark contrast to this assumption, the neurobiological data reviewed here suggest that much, if not all, of neocortex is multisensory… The world is [a] barrage of sensory inputs, our perception is a unified representation of it, and the neocortex is organized in a manner to make the underlying processes as efficient as possible.

Certainly this makes excellent sense of many aspects of musical experience. Music maps onto so much else, in this case, not only because sounds are not tied to particular semantic meanings but also, crucially, because so much of the neocortex can do something with it. The key feature is change in experience over time, tied (via the sensorimotor cortex) to a sense of movement. That links together changes in space with changes in sensory and emotional experience, just as we find in the case of musical shape. But shape is only one instance, albeit a particularly general and semantically flexible one, of a connection that a multisensory capacity in the sensorimotor cortex would facilitate: we easily find in music the characteristics of many very specific kinds of motions (stumbling, plodding, jerking, drooping and so on) or behaviours (clumsy, tentative, alluring, rapturous, imploring), all combining aspects of movement with sensation and narrative implication (again we have action, imaginative simulation and understanding).

So the sensorimotor system begins to look like a metaphor engine for anything that involves an experience of motion in relation to imagination and meaning. It receives input from the other modalities and conceives it (models it) in terms of dynamics with potential for meanings. It generates a 'vitality form' in Stern's sense, or a shape in ours, a core experience with dynamic characteristics independent of any particular modality; and the same form or shape can easily be perceived as like the shape or form of experiences in other modalities. Here metaphor becomes embodied, or (perhaps better) the embodied becomes metaphor.

In conclusion

The kinds of processes that are being revealed by research in the multisensory brain go a long way towards explaining how sound could so easily be mapped onto shape and why it might seem so 'natural' to musicians (and especially western musicians) to think of musical performance as giving shape to the notes in the score. It is an immensely flexible concept, very easy, because it involves changing quantity and intensity, to apply in any aspect of sound and experience where it can do some useful work. In performance, as in other kinds of experience, shape can apply to anything that changes. In shaping a score, performers can select one or more of many available dimensions in sound. And this multi-applicability of shape to sound enables them to be responsive to context, shaping different dimensions from moment to moment. It enables them to define a personal approach to shaping as part of their own performance style, identifying them, communicating their way of understanding music. It enables them to work personally within a constrained period style, itself defined by particular ways of shaping sound. But this same flexibility also affords the possibility of massive change in period style over time (Leech-Wilkinson 2009b). This is exactly what we find now that we have well over one hundred years of recorded performance.

The fundamental level on which a sense of shape can be passed around the brain could be a crucial factor in allowing music to be made expressive by performers in a shifting variety of complex and interesting ways, and it explains how the concept of shape can be used to think about and act within all of them. Shape gets us about as close as we can get to suprasensory modalities. It works in every dimension that can change over time. And this is why it is so useful to performers in thinking and talking about how to make music. It is one of the most powerful ways we have of making sense of the experience of music without having to be too specific.

A performer in a recent interview for Prior (2011) said: 'And that's the thing about music, if you use imagery it makes your muscles do all kinds of things that you don't necessarily have to describe in a minutely physical way.' This is a crucial point. Instead of having to say, 'I want you to make the A in bar 3 slightly softer and maybe 40 milliseconds longer than the previous G, and then the B semiquaver a little bit longer, maybe another 20 ms, and about 20 dB quieter than you'd expect, *or* alternatively you could make the A slightly louder and shorter and the G very short, *or* . . .' and so on, you can simply say, 'I'd like you to shape that phrase a little more', and then the performer does whatever s/he feels works. Shape is great value for performers, communicating the effect that is required, leaving them free to produce it through feeling, not analysis (Leech-Wilkinson and Prior 2014; perhaps assisted by auditory imagery: Keller, Dalla Bella and Koch 2010), and thus, free to use their experience, judgement and taste, to express their musicianship. And we as listeners align ourselves to the shaped template that the musician supplies (DeNora 2004). Performed shapes, arranged in a persuasively and movingly managed sequence, become our felt experience (Johnson 2007: 238). I suggested above that music is lifelike in a utopian fantasy world where everything that happens makes sense and where, whatever conflicts we may experience along the way, everything turns out for the best. I was really speaking of composition then. But we can say something very similar about performance and the way it uses the notion of shape: to speak of shape in performance is to speak of the way music, given a highly skilled performer, enacts an idealized image of the feeling experience of our everyday lives.

In sum, we have seen that shape is a highly flexible concept widely used by (especially, but not only) western musicians (Prior, Chapter 7) to talk about the expressive qualities of a performance. It relates closely to other concepts involving real or imagined motion through space (including gesture and trajectory) or across terrain (landscape, contour). At a more general level it conceptualizes change over time. But fundamentally, in all this discourse, shape is modelling changing feelings, and it is that mapping between the dynamics of musical sound and the dynamics of feelings that allows shape to function so effectively as a way of thinking and speaking about musical expressivity. The dynamics of musical sound are easily analysed and visualized with sound visualization and mapping software (such as Sonic Visualiser) which gives some access to the shaped nature of performance and the expressive work it does for the listener. But the underlying mechanisms need to be teased out through other kinds of research. Stern's 'forms of vitality' offer a powerful means of thinking further about the shape of feeling, while Johnson's work on embodiment and image schemas helps to show how the relationship between shape and feeling is grounded in bodily experience. Research suggesting the existence of modes of suprasensory perception can be linked to work on the neural mechanisms of sensory perception that arrives at similar conclusions. Research on multimodal perception offers a neural mechanism by which a sense of shape may be generated simultaneously in sound, vision and motion, and sheds additional light on the ease with which music is described using metaphor and in terms of its likeness to other things, the means by which it acquires so many of the meanings attributed to it.

Bringing this work together under the umbrella of shape brings us closer to understanding both how a sense of shape is generated by musical sound and how and why it is so effective as a tool for musicians. For the end user, thought about and talk of shape enables the defining and sharing of ideas about an activity (expressive performance) largely carried out within the domain of feeling and intuitive response, a domain otherwise inaccessible to analysis or discussion. Shape functions, then, on at least two levels. As a way of talking generally about the dynamics of performance, it affords efficient and communicative teaching and rehearsing. As a model of the dynamics of performance, it encapsulates the changing magnitudes of sound during performance, affording a sense of contoured trajectory through which feeling and sound can be aligned.

References

Aksnes, H., 2001: 'Music and its resonating body', *Danish Yearbook of Musicology* 29: 81–100.

Altenmüller, E., M. Wiesendanger and J. Kesselring, 2006: *Music, Motor Control and the Brain* (Oxford: Oxford University Press).

Antovic, M., 2009: 'Musical metaphors in Serbian and Romani children: an empirical study', *Metaphor and Symbol* 24/3: 184–202.

Athanasopoulos, G. and N. Moran, 2013: 'Cross-cultural representations of musical shape', *Empirical Musicology Review* 8/3–4: 185–99.

Bueti, D. and V. Walsh, 2009: 'The parietal cortex and the representation of time, space, number and other magnitudes', *Philosophical Transactions of the Royal Society B* 364: 1831–40.

Currie, G., 2011: 'Empathy for objects', in A. Coplan and P. Goldie, eds., *Empathy: Philosophical and Psychological Perspectives* (Oxford: Oxford University Press), pp. 82–97.

DeNora, T., 2004: 'Historical perspectives in music sociology', *Poetics* 32: 211–21.

Doğantan-Dack, M., 2013: 'Tonality: the shape of affect', *Empirical Musicology Review* 8/3–4: 208–18.

Dolscheid, S., S. Shayan, A. Majid and D. Casasanto, 2013: 'The thickness of musical pitch: psychophysical evidence for linguistic relativity', *Psychological Science* 24/5: 613–21.

Eitan, Z., 2013: 'Musical objects, cross-domain correspondences, and cultural choice: commentary on "Cross-cultural representations of musical shape" by George Athanasopoulos and Nikki Moran', *Empirical Musicology Review* 8/3–4: 204–7.

Eitan, Z. and R. Y. Granot, 2006: 'How music moves: musical parameters and listeners' images of motion', *Music Perception* 23/3: 221–47.

Eitan, Z. and R. Y. Granot, 2007: 'Intensity changes and perceived similarity: interparametric analogies', *Musicae Scientiae, Discussion Forum* 4A: 39–75.

Eitan, Z. and R. Timmers, 2010: 'Beethoven's last piano sonata and those who follow crocodiles: cross-domain mappings of auditory pitch in a musical context', *Cognition* 114: 405–22.

Fraisse, P., 1984: 'Perception and estimation of time', *Annual Review of Psychology* 35: 1–36.

Gallese, V. and G. Lakoff, 2005: 'The brain's concepts: the role of the sensory-motor system in conceptual knowledge', *Cognitive Neuropsychology* 22/3: 455–79.

Ghazanfar, A. A. and C. E. Schroeder, 2006: 'Is neocortex essentially multisensory?', *Trends in Cognitive Sciences* 10/6: 278–85.

Hausegger, F. v., 1887: *Die Musik als Ausdruck* (Vienna: Carl Konegen).

Huron, D., 2006: *Sweet Anticipation: Music and the Psychology of Expectation* (Cambridge, MA: MIT Press).

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

Johnson, M. L. and S. Larson, 2003: '"Something in the way she moves": metaphors of musical motion', *Metaphor and Symbol* 18: 63–84.

Juslin, P. N. and E. Lindström, 2010: 'Musical expression of emotions: modelling listeners' judgements of composed and performed features', *Music Analysis* 29/1–3: 334–64.

Keller, P. E., S. Dalla Bella and I. Koch, 2010: 'Auditory imagery shapes, movement timing, and kinematics: evidence from a musical task', *Journal of Experimental Psychology: Human Perception and Performance* 36: 508–13.

Kim, J. H., 2012: 'What music and dance share: dynamic forms of movement and action-based aesthetic empathy,' in S. Schroedter, ed., *Bewegungen zwischen Hören und Sehen Denkbewegungen zu Bewegungskünsten* (Würzburg: Königshausen & Neumann).

Kim, J. H., 2013: 'Shaping and co-shaping forms of vitality in music: beyond cognitivist and emotivist approaches to musical expressiveness', *Empirical Musicology Review* 8/3–4: 162–73.

Koelsch, S., 2010: 'Towards a neural basis of music-evoked emotions', *Trends in Cognitive Sciences* 14/3: 131–7.

Köhler, W., 1929: *Gestalt Psychology* (New York: Liveright).

Küssner, M. B. and D. Leech-Wilkinson, 2014: 'Investigating the influence of musical training on cross-modal correspondences and sensorimotor skills in a real-time drawing paradigm', *Psychology of Music* 42/3: 448–69.

Küssner, M. B., D. Tidhar, H. M. Prior and D. Leech-Wilkinson, 2014: 'Musicians are more consistent: gestural cross-modal mappings of pitch, loudness, and tempo in real-time', *Frontiers in Psychology* 5/789, doi: 10.3389/fpsyg.2014.00789 (accessed 9 April 2017).

Lakoff, G. and M. Johnson, 1980: *Metaphors We Live By* (Chicago: Chicago University Press).

Langer, S. K., 1963: *Philosophy in a New Key*, 3rd edn (Cambridge, MA: Harvard University Press).

Lavy, M. M., 2001: 'Emotion and the experience of listening to music: a framework for empirical research' (PhD dissertation, University of Cambridge).

Leech-Wilkinson, D., 2006: 'Expressive gestures in Schubert singing on record', *Nordisk Estetisk Tidskrift* 33–4: 50–70.

Leech-Wilkinson, D., 2009a: *The Changing Sound of Music: Approaches to the Study of Recorded Musical Performances*, http://www.charm.kcl.ac.uk/studies/chapters/intro.html (accessed 9 April 2017).

Leech-Wilkinson, D., 2009b: 'Recordings and histories of performance style', in N. Cook, E. Clarke, D. Leech-Wilkinson and J. Rink, eds., *The Cambridge Companion to Recorded Music* (Cambridge: Cambridge University Press), pp. 246–62.

Leech-Wilkinson, D. and H. Prior, 2014: 'Heuristics for expressive performance', in D. Fabian, R. Timmers and E. Schubert, eds., *Expressiveness in Music Performance: Empirical Approaches Across Styles and Cultures* (Oxford: Oxford University Press), pp. 34–57.

Levitas, R., 2010: '*In eine bess're Welt entrückt*: reflections on music and utopia', *Utopia Studies* 21/2: 215–21.

Lipps, T., 1903: 'Einfühlung, innere Nachahmung, und Organempfindungen', *Archiv für die gesamte Psychologie* 1: 185–204.

Ludwig, V. U., I. Adachi and T. Matsuzawa, 2011: 'Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans', *Proceedings of the National Academy of Sciences* 108/51: 20661–5.

Marks, L. E., 1978: *The Unity of the Senses: Interrelations among the Modalities* (New York: Academic Press).

Martino, G. and L. E. Marks, 2000: 'Cross-modal interaction between vision and touch: the role of synesthetic correspondence', *Perception* 29: 745–54.

Maurer, D. and C. J. Mondloch, 2006: 'The infant as synesthete', *Attention and Performance* 21: 449–71.

Maurer, D., T. Pathman and C. J. Mondloch, 2006: 'The shape of boubas: sound–shape correspondences in toddlers and adults', *Developmental Science* 9/3: 316–22.

McLachlan, N. and S. Wilson, 2010: 'The central role of recognition in auditory perception: a neurobiological model', *Psychological Review* 117/1: 175–96.

McLachlan, N. M., L. J. Greco, E. C. Toner and S. J. Wilson, 2011: 'Using spatial manipulation to examine interactions between visual and auditory encoding of pitch and time', *Frontiers in Psychology* 1/233, doi 10.3389/fpsyg.2010.00233 (accessed 9 April 2017).

Miller, G., 2000: 'Evolution of human music through sexual selection', in N. L. Wallin, B. Merker and S. Brown, eds., *The Origins of Music* (Cambridge, MA: MIT Press), pp. 329–60.

Molnar-Szakacs, I. and K. Overy, 2006: 'Music and mirror neurons: from motion to "e"motion', *Social Cognitive and Affective Neuroscience* 1: 235–41.

Murray, M., 2003: 'Narrative psychology', in J. A. Smith, ed., *Qualitative Psychology: A Practical Guide to Research Methods* (London: Sage), pp. 111–31.

Näätänen, R. and I. Winkler, 1999: 'The concept of auditory stimulus representation in cognitive neuroscience', *Psychological Bulletin* 125: 826–59.

Nakamura, J. and M. Csikszentmihalyi, 2009: 'Flow theory and research', in C. R. Snyder and S. J. Lopez, eds., *Oxford Handbook of Positive Psychology* (Oxford: Oxford University Press), pp. 195–206.

Narmour, E., 1990: *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model* (Chicago: University of Chicago Press).

Peretz, I., 2003: 'Brain specialization for music: new evidence from congenital amusia', in I. Peretz and R. Zatorre, eds., *The Cognitive Neuroscience of Music* (Oxford: Oxford University Press), pp. 192–203.

Prince, J. B., M. A. Schmuckler and W. F. Thompson, 2009: 'Cross-modal melodic contour similarity', *Canadian Acoustics* 37/1: 35–49.

Prior, H. M., 2010: 'Links between music and shape: style-specific; language-specific; or universal?', paper presented at 'Topics in Musical Universals: 1st International Colloquium', Aix-en-Provence, France, December 2010.

Prior, H. M., 2011: 'Exploring the experience of shaping music in performance', paper presented at the Performance Studies Network First International Conference, Cambridge, UK, 14–17 July 2011.

Prior, H. M., 2012: 'Shaping music in performance: report for questionnaire participants (revised August 2012)', http://www.cmpcp.ac.uk/wp-content/uploads/2015/09/Prior_Report.pdf (accessed 9 April 2017).

Ramachandran, V. S. and E. M. Hubbard, 2001: 'Synaesthesia: a window into perception, thought and language', *Journal of Consciousness Studies* 8/12: 3–34.

Rees, A. and A. R. Palmer, eds., 2010: *The Oxford Handbook of Auditory Science: The Auditory Brain* (Oxford: Oxford University Press).

Spence, C., 2011 'Crossmodal correspondences: a tutorial review', *Attention, Perception, and Psychophysics* 73/4: 971–95.

Spence, C. and O. Deroy, 2012: 'Crossmodal correspondences: innate or learned?', *i-Perception* 3: 316–18.

Stein, B. E. and T. R. Stanford, 2008: 'Multisensory integration: current issues from the perspective of the single neuron', *Nature Reviews: Neuroscience* 9: 255–66 and corrigendum.

Stern, D., 2004: *The Present Moment in Psychotherapy and Everyday Life* (New York: Norton).

Stern, D., 2010: *Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy, and Development* (Oxford: Oxford University Press).

Stevens, C. J., E. Schubert, R. Haszard Morris, M. Frear, J. Chen, S. Healey, C. Schoknecht and S. Hansen, 2009: 'Cognition and the temporal arts: investigating audience response to dance using PDAs that record continuous data during live performance', *International Journal of Human-Computer Studies* 67: 800–13.

Tarasti, E., 2004: 'Music as a narrative art', in M.-L. Ryan, ed., *Narrative across Media: The Languages of Storytelling* (Lincoln: University of Nebraska Press), pp. 283–304.

Truslit, A., 1938: *Gestaltung und Bewegung in der Musik* (Berlin: Chr. Friedrich Vieweg).

Walker, P., J. G. Bremner, U. Mason, J. Spring, K. Mattock, A. Slater and S. P. Johnson, 2010: 'Preverbal infants' sensitivity to synaesthetic cross-modality correspondences', *Psychological Science* 21: 21–5.

Walsh, V., 2003: 'A theory of magnitude: common cortical metrics of time, space and quantity', *Trends in Cognitive Sciences* 7: 483–8.

Ward, J., B. Huckstep and E. Tsakanikos, 2006: 'Sound-colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all?', *Cortex* 42: 264–80.

Ward, J., S. Moore, D. Thompson-Lake, S. Salih and B. Beck, 2008: 'The aesthetic appeal of auditory–visual synaesthetic perceptions in people without synaesthesia', *Perception* 37: 1285–96.

Wingstedt, J., S. Brändström and J. Berg, 2010: 'Narrative music, visuals and meaning in film', *Visual Communication* 9: 193–210.

Yu, N., 2008: 'Metaphor from body and culture', in R. W. Gibbs, ed., *The Cambridge Handbook of Metaphor and Thought* (Cambridge: Cambridge University Press), pp. 247–61.