

Joel Krueger

Enacting Musical Content

According to the enactive view, perception is intrinsically active (Hurley 2001; Noë 2004; O'Regan and Noë 2001; Thompson 2005). Perceptual experience acquires content – objects of experience are brought to phenomenal presence – via the possession and deployment of bodily skills: the ability to use our body to explore, manipulate, and engage with the world and things in it. Accordingly, enactive approaches tend to be pitched as theories of *access* (e.g. Noë 2009): accounts of how the possession, deployment, and understanding of bodily skills, as well as the sensorimotor regularities governing our engagement with the world, determine the character (the *how*) and content (the *what*) of perceptual experience – that is, the form of our experiential access to the world. Enactive views take as their target classical computational theories of vision which, according to the enactivist, fail to account for the crucial contribution that agency makes to visual experience (Gangopadhyay and Kiverstein 2009). Though not without their critics, enactive approaches are increasingly influential. Yet proponents have said little of how the account might be extended to sensory modalities other than vision and touch.

This chapter offers the beginning of an enactive account of auditory experience – particularly the experience of listening sensitively to music. It investigates how sensorimotor regularities grant perceptual access to music *qua* music. Two specific claims are defended: (1) music manifests experientially as having complex spatial content; (2) sensorimotor regularities constrain this content. Musical content is thus brought to phenomenal presence by bodily exploring structural features of music. We enact musical content.

1 Preliminaries

Why is this discussion philosophically interesting? First, this view extends enactive accounts of perception beyond the realm of vision and touch. Most discussions of perception within enactive literature take vision as the paradigm case of perceiving. Yet the core claims animating the enactive

view apply equally well to auditory perception, I suggest, including the experience of listening to music. Moreover, they allow for an adequate phenomenological portrayal of our sensitive listening episodes – instances of what I will term ‘deep listening’ – one which captures the experiential features unique to musical experience. Finally, these enactive claims receive robust empirical support from several strands of empirical studies on auditory and music perception, discussed below.

Second, the enactive view here developed challenges dominant tendency in philosophy and psychology to treat musical experience as a relatively passive affair – that is, as consisting of a linear causal process leading from musical piece to listener. John Sloboda has labelled this the ‘pharmaceutical model’ of musical experience (Sloboda 2005). According to Sloboda, this model rests on the assumption that music’s psychological efficacy is much like taking medication. There are pre-determined, consistent effects that result from specific musical structures interacting with specific brain regions (*ibid.*, p. 319). But like Sloboda, I will argue that this ‘passive listener model’ of music listening ignores the richly interactive nature of musical experience. Musical experience is a form of active perceptual exploration – an active, world-engaged *musicking* (Small 1998).¹ The enactive view here defended thus emphasizes the dynamic and agentive nature of music perception, urging that the embodied and situated listener has a central role in shaping both the character and content of musical experience. Musical experience, like perceptual consciousness more generally, is *transactional* (Putnam 1999).

2 The space of hearing

Do we auditorily experience spatial features of sounds? Brian O’Shaughnessy states flatly that ‘we absolutely never perceive sounds to be *at* any place. (Inference from auditory data being another thing.)’ (O’Shaughnessy 1984, p. 199). This is because sounds have *aspatial* phenomenology. O’Shaughnessy doesn’t deny, of course, that we can *locate* sounds within the surrounding environment. However, this locational aspect of sound perception is determined by *acoustic* features of the sounds: for example, pitch, timbre, or loudness, such as recognizing that the rising loudness of an ambulance’s siren means that it is spatially closer to me than it was ten seconds ago. But again, there is no spatial content intrinsic to the phenomenology of siren hearing on its own. All we hear, rather, is the sound itself. We don’t hear the sound as standing in any relation to the space it occupies

1 F. Joseph Smith observes similarly that, ‘[t]he image of the extended hand, significative of intentionality, connotes the activity required to constitute the proper receptivity. Receiving is, therefore, not inert passivity. The hand must close on the gift, else it will fall to the ground. Similarly, in order to listen to either spoken language or to musical sound, we have to “lend our ears.” Listening is not merely passive reaction; it is its own activity’ (Smith 1979, pp. 100–111).

(Nudds 2001, pp. 213–4). This is because ‘the sound I hear is where I am when I hear it’ (O’Shaughnessy 1984, p. 199). O’Shaughnessy thus endorses a ‘proximal’ theory of sounds according to which sounds are located where their hearer is (Casati and Dokic 2009). When we locate sounds in space, the auditory data present to our ears is ‘augmented by mental factors leading to one’s hearing the sounds to be *coming from* a specific site’ (O’Shaughnessy 2009, p. 125). We infer or work out – again, via acoustic features of sounds or by extra-auditory data from experiences in other modalities (e.g. vision) that do possess genuine spatial content – where the sound sources are likely located. But the sounds-as-heard have no intrinsic spatial phenomenology. Rather, audition plays a supplementary role. According to O’Shaughnessy, it provides inference cues to spatial features and locations that come to us through other sense modalities. However, lacking non-derived spatial content, audition in this way stands in contrast to experiences of other sense modalities.

There are some problems with this view salient to this paper’s core concerns. First, a simple phenomenological objection is that the characterization of hearing at work here misdescribes the experiential character of everyday auditory experience. For, we surely don’t have the extra-auditory *experience* of inferring or working out the direction and location of sounds. Rather, we hear the sounds themselves, immediately and non-inferentially, as egocentrically located in our surrounding environment. For instance, I immediately hear the laughter and voices of my neighbour’s children playing outside *slightly behind me and to the left*, just outside the window of my office; and if I suddenly hear a dull thud followed by crying – perhaps one of the children has run into the front yard and tripped over the sprinkler along the way – I know exactly where to go to offer aid. Likewise, I hear the bird squawking loudly as he passes by *directly above me*; the voices of the couple arguing emanate from the apartment *right next to mine*. When someone calls my name, I immediately turn without having to think about it in the direction of the sound; and blind people, too, are quite capable of tracking and responding to sound events despite the absence of extra-auditory visual input (Hamilton 2009, p. 176). Phenomenologically speaking, it appears that auditory experiences are locational. They represent both *what* is happening (e.g. children playing outside) as well as *how* what is happening stands in relation to oneself (e.g. slightly behind me and to the left).

Moreover, our ability to non-inferentially locate sounds in egocentric space purely by hearing them has important behavioural consequences. If, while attending a baseball game and chatting up a friend in the stands, I suddenly hear the crack of a bat and the whizz of an oncoming baseball, I know immediately which way to duck without having to first consider extra-auditory data (which in this case is likely not even available to me,

given that I'm looking away from the sound source). The ability to track and quickly respond to events in the environment rests on having immediate phenomenal access to the spatial location carried by sounds. A step-wise process depending upon the access and utilization of extra-auditory data, or necessarily mediated by an inferential 'working out' of a sound's location, places an unnecessarily excessive computational burden on the perceiver – and thus would, accordingly, significantly impede their reaction times. O'Shaughnessy's view is thus phenomenologically implausible (it is also challenged by empirical research, as will become apparent throughout this paper). And while this phenomenological objection is not in itself a devastating refutation, the view's descriptive implausibility should at least give us some pause.

Another difficulty with the view, as Casey O'Callaghan notes, is the paucity of egocentric information provided by purely *acoustic* features of sounds, such as pitch, timbre, and loudness (O'Callaghan 2010, p. 133). According to O'Shaughnessy, acoustic features of sounds are the primary bearers of whatever spatial data is available from sounds. Again, we hear the rising loudness of a siren and work out from this loudness that it is moving closer to us. But acoustic features seem ill-suited to fill this role. For 'perceptible qualitative attributes, such as pitch, timbre, and loudness fail to correspond reliably to egocentric location, and variations in qualities do not correspond reliably to changes in egocentric location' (O'Callaghan 2009, p. 133). So, pitch and timbre preserve their character independently of spatial location: a musical instrument may exhibit the same pitch and timbre whether it's right next to the hearer or across the street. Similarly, while loudness is perhaps a more reliable indicator of location (e.g. the loud voice on the train is coming from the man in the seat immediately behind mine), there are no fixed sensorimotor rules governing the relation between loudness and distance. A musical instrument playing softly may be right next to the hearer while a loudly playing one is across the street. Some researchers have suggested that loudness constancy may not depend upon, or even be related to, source distance perception at all. Zahorik and Wightman conducted several experiments in which listeners reported robust loudness constancy even when the source distance was varied (Zahorik and Wightman 2001). Listeners also systematically overestimated close-range sources and underestimated long-range source distances (*ibid.*, p. 81). Given the facility with which we track and respond to sound-events in our environment – such as the baseball example discussed above – it is therefore unlikely that this facility turns on our responsiveness to purely acoustic features which, as we've just seen, have a relatively tenuous correspondence relation to the listener's egocentric location. It is rather more likely that sounds themselves bear spatial content.

Finally, one more line of empirically-informed support for this idea comes from work on the neural representation of ‘auditory space maps’: a neural map of how received auditory information is situated in the surrounding environment (Hyde and Knudsen 2002). It appears that ‘[a]uditory space maps can be generated without visual input, but their precision and topography depend on visual experience. So, for example, owls raised as if they were blind end up with abnormal, or even partially inverted, auditory maps’ (Carr 2002, p. 30). While visual input provides more reliable and topographically organized information—and thus can refine and enhance the auditory information represented by auditory space maps—evidence nevertheless suggests that spatial content cuts across both visual and auditory perception (not to mention other modalities). It is thus a mistake to think of auditory experience as bearing only derived spatial content.

In sum: listening to sounds is an exploration of our world—including spatial and locational aspects of things in it. Sounds routinely furnish spatial information about our world, and we use our auditory experiences to explore and skillfully respond to things happenings in it. Phenomenologically, sounds are thus spatially structured.

3 Exploring musical space

One of the things we quite often hear in our world is music, both live and recorded. I now want to argue that the phenomenology of musical experience is determined by the experience of two forms of spatiality: what I term, respectively, ‘inner’ and ‘outer’ musical space. I will argue that, in episodes of ‘deep listening’—listening in a voluntary mode of sustained perceptual focus and affective sensitivity, as opposed to hearing music with ‘one ear’ as a piece drifts idly by in the background—listeners enact the experiential fusing of these two forms of spatiality.² Put otherwise, the spatiality of musical structure—and in particular, structural features like *textural qualities* and the temporal regularities of *sonic patterns* (both melodic and rhythmic)—presents music as having an exploratory profile affording this sort of deep listening. My approach in this section and the next one is to offer a descriptive phenomenology of how we enact the spatial content of the form of musical experience I am calling ‘deep listening’. The section thereafter then discusses some empirical research that seems to support this phenomenological description.

Experience, according to the enactivist, is always an active encounter with hidden complexity (Noë 2009, p. 473). Conscious phenomena harbour potentially attended-to aspects that invite further exploration. In this

² Composer Pauline Oliveros coined the expression ‘deep listening’ to refer to her own brand of music composition and listening practices. See <http://deeplisting.org>. See also Becker (2004).

case of visual perception, the anticipation of how a visually conscious object changes relative to bodily movements disclosing previously-hidden aspects of the object (e.g. leaning to the left or right to bring unseen bits of a solid object into view) is a crucial part of actively perceiving the world. These anticipations or expectations are a form of sensorimotor knowledge—an understanding of how our perceptual relation to the world is mediated by contingent relations coupling bodily movement and sensory change. As noted above, most enactive literature focuses on visual perception as the paradigm case of enacting experiential content. But thinking of experience as an active encounter with hidden complexity is no less true for audition and music perception than it is for other forms of experience. Music, in particular, invites sensitive perceptual inspection. It solicits inspection of discrete constituent units that can be individually attended to, differentiated, and sonically explored.³ This is because musical structure consists of sonic units extended in time—unlike units of visual objects, which are extended in space—and which are therefore perceptually individuated in virtue of pitch and other temporal characteristics (O’Callaghan 2009). This dynamic temporal structure presents an especially rich sort of exploratory profile.

Temporality is essential both to the structure of musical experience as well as to the exploratory profile a musical piece presents. As Søren Kierkegaard notes, ‘aside from language, music is the only medium that addresses itself to the ear... Language has time as its element; all other media have space as their element. Music is the only other one that takes place in time’ (Kierkegaard 1959, pp. 66, 67).⁴ Of course, other media are also situated in time and can, accordingly, exhibit various changes that betray this temporality. Colours on a painting gradually fade as it ages; shadows pass across the surface of a sculpture, giving it distinctive appearances in the light of early morning versus the dim hues of evening. Kierkegaard’s point, however, is that with music—unique among the arts—‘temporality is not a matter of “subjectivity” but a matter of the way the phenomenon presents itself’ (Ihde 2007, p. 94). I can perceptually explore a painting or a statue—I can sit and gaze intently; walk up to it, move away, tilt my head and look from another angle; touch it and run my hand over its surface, etc. — and in this way become aware of its nature as a temporally situated object. But this movement marks a shift toward the

- 3 Marilyn Nonken observes that, ‘[d]esigned intentionally for sensory exploration, the musical environment is characterized by the presence of not only harmony and rhythm but also such factors as silence, timbre (instrumentation), dynamic (amplitude), density, texture, gestural and motivic figures, patterns, and audible processes of accretion and degradation (such as *crescendo* or *ritardandi*, the processes of getting louder or slower)’ (Nonken 2008, p. 294).
- 4 Schopenhauer puts the point more strongly when he insists that music is perceived ‘in and through time alone, with the absolute exclusion of space’ (Schopenhauer 1966, p. 266). It is precisely this view that I will challenge, arguing that both temporality and spatiality are integral parts of music and musical experience.

noetic phenomena—that is, my consciousness *as aware* of the passing of time within my exploratory activity (Ihde 2007, p. 94). Temporality is not immediately manifest within the object *itself* (e.g. the painting as *noematic* correlate) but rather within my intentional relation *to* the object. On the other hand, there seems to be a unique structural intimacy between music and temporality; the latter is an essential part of the former's make-up. We therefore cannot hear music without simultaneously hearing how time is embodied within the music.

Yet by focusing exclusively on the temporality of musical experience, there is a danger of losing phenomenological grip on its inherently spatial qualities. A central feature of music's exploratory profile is space. For, as Robert Morgan notes, 'it would seem to be impossible to talk about music at all without invoking spatial notions of one kind or another' (Morgan 1980, p. 527).⁵ To perceive music is thus to perceive space. Though I argued above that all auditory experience bears spatial content, musical experience is unique, I suggest, in terms of the *complexity* of its spatial content. Moreover, with respect to other forms of art, the space of music is perceived precisely *in* its temporality in a way not the case with other art forms.

To begin to get a sense of how this is so, consider first some apt remarks by Merleau-Ponty. Despite a relative lack of interest in music in his writing, Merleau-Ponty nevertheless offers a few quotes of interest. First, like Kierkegaard, he emphasizes the importance of temporality in perceiving music. A piece of music, he says,

...comes very close to being no more than a medley of sound sensations: from among these sounds we discern the appearance of a phrase and, as phrase follows phrase, a whole and, finally, as Proust put it, a world. This world exists in the universe of possible music, whether in the district of Debussy or the kingdom of Bach. (Merleau-Ponty 2004, p. 99)

The implication seems to be that the 'world-making' power of music only becomes apparent through the active exploration of a piece—that is, careful attentiveness to the temporal dynamics of a piece of music (i.e. 'as phrase follows phrase') that gradually erect a sonic topography inviting further exploration. The temporal unfolding of music is the movement that begins to open up a piece's inner sonic space (I clarify what I mean by this in a moment).

Elsewhere, however, we find a more substantive phenomenological observation:

⁵ Morgan writes further that, 'Thus in discussing even the most elementary aspects of pitch organization—and among the musical elements, only pitch, we should remember, is uniquely musical—one finds it necessary to rely upon such spatially oriented oppositions as "up and down," "high and low," "small and large," (in regard to intervallic "distances"), and so on' (Morgan 1980, p. 527).

When, in the concert hall, I open my eyes, and visible space seems to me cramped compared to that other space through which, a moment ago, the music was being unfolded, and even if I keep my eyes open as the piece is being played, I have the impression that the music is not really contained within this circumscribed and unimpressive space. It brings a new dimension stealing through visible space, and in this it surges forward... (Merleau-Ponty 2002, pp. 257–258)

There are several points of interest in this short passage. Salient to present concerns is the claim that musical experience, while temporal, is additionally infused with representations of space. Music both consumes as well as creates space. Specifically, we can say that Merleau-Ponty differentiates between what we might term ‘inner’ and ‘outer’ musical space. The former refers to the space internal to the piece of music itself. It is what we might term *structural* space: that is, the piece’s inner syntactical structure established by the way that constituent components (e.g. tones, rhythmic progressions, etc.) hang together, lending the musical piece its sonic coherence as a composed object. This form of musical space is experientially fluid; it can swell and expand, as when a piece of music seems to fill a room and surround us, occupying ‘a new dimension stealing through visible space’. ‘Outer’ musical space, on the other hand – what Merleau-Ponty calls ‘visible space’ – might also be termed *locational* space. This is the egocentric spatial character of music as locationally perceived (e.g. music heard as coming from the speakers *in front of me* or the stage *to my right*), as something inhabiting a determinate location relative to my bodily orientation.

When we perceive a piece of music, we tend to automatically perceive the piece’s inner spatial configuration. This is what it means to listen to music understandingly, to hear it as something with an inner complexity offering up an exploratory profile inviting attentive inspection.⁶ And we also tend to have a reasonably clear sense of where the musical source is located spatially, such as when we walk into an unfamiliar apartment for the first time and immediately recognize that the stereo is playing in the next room. However, what I want to suggest is that within deep listening episodes, we enact an experiential fusing of these two forms of musical spatiality such that neither takes phenomenological precedence over the other. Rather, they come together and, in their fusing, open up experiential character of the piece in a new and previously unheard way. This is what gives these episodes their unique phenomenal character. And the animate body, as we will see, plays a central role in facilitating this musical-spatial enaction.

Again, ‘deep listening’, as I’m using this expression, is a voluntary form of musical experience consisting of sustained attentional focus and affec-

⁶ The phenomenological consequences of failing to perceive the inner space of music will be discussed in more detail below.

tive sensitivity. It is an immersive form of listening in which the subject *selectively orients* herself to a piece of music by actively attending to its various sound features and their interrelationships – while simultaneously maintaining a state of affective receptivity, or a readiness-to-be-moved, by what is happening sonically in the music. Deep listening is thus a transactive mode of listening involving ‘processes such as *exploring, selecting, modifying, and focusing of attention*’ (Reybrouck 2005, p. 252). Moreover, this deep engagement can have the temporary effect of weakening or obliterating the felt senses of inner and outer. This is the cultivation of an auditory *field state*: an expanded, phenomenally ‘full’ mode of listening in which focal attention stretches to the very boundaries of the sound as present (Ihde 2007, p. 102). In other words, instead of remaining remotely situated, the deep listener instead has the felt sense of *inhabiting* the sound field, leading to a heightened emotional and affective responsiveness to the musical situation (Vastfjall 2003).

Deep listening is thus sensually richer than involuntary or passive modes of hearing such as being faintly aware of background music playing in a grocery store or restaurant, or hearing the sound of a radio drifting out of a nearby open window. In these latter cases, the *locational* spatial character of a musical piece remains experientially prioritized. For example, we hear unobtrusive *muzak* trickle quietly from speakers above us, staggered across the ceiling of the grocery store; we cringe slightly at the shrill sound of a teenager in the bus seat behind ours listening to hip-hop via the underpowered speaker of his mobile phone; the sound of a radio is momentarily present before slowly diminishing and trailing off as we walk by an open window in a nearby apartment complex. However, since our attention is largely focused elsewhere within these shallow listening episodes (e.g. navigating our shopping cart toward the exit; peering out the window to see if our stop is coming up; hurrying on to make the appointment for which we’re already late), the inner or structural space of the piece fails to present itself within any sort of phenomenal immediacy. It is experientially present – again, to perceive a particular auditory event *qua* music is to perceive its inner space, however dimly – but it remains diminished and nonfocal, confined instead to the relative margins of our awareness. This passive hearing is thus a minimally active form of musical engagement due to our lack of attentive inspection and affective engagement. Once more, the outer spatiality of the piece is given phenomenological priority within this mode of hearing; it remains predominantly allocentric in character (i.e. ‘The sound of the radio is *over there*’ – Turner *et al.* 2007).⁷

7 This is not to suggest that there isn’t an egocentric aspect to allocentric hearing. For the ‘over there’ of heard music is the disclosure of an ‘over there’ specified as such only in relation to my bodily *here*. So, allocentric and egocentric information within audition are experientially

4 Solitary and social forms of deep listening

I've suggested that the experiential character, as well as the representation of the relation between inner and outer musical space, differs in deep listening episodes. While perhaps a somewhat rarified occurrence, deep listening can nevertheless be enacted within nearly any context where music is carefully attended to (and, indeed, the auditory conditions are sufficiently adequate). It can be a solitary undertaking or, in some contexts, a social affair (more on that below). Importantly, it is an intentional modification of everyday hearing—a kind of playing with perception, so to speak—that emphasizes the unique agility and, indeed, *plasticity* of audition. Don Ihde has helpfully observed that the spatiality of the auditory field exhibits a 'double dimensionality': it simultaneously exhibits both *surroundability* (i.e. an atmospheric or enveloping quality) and *directionality* (i.e. a situatedness or locality) (Ihde 2007, p. 77). This double dimensionality is both a source of modality-specific ambiguity as well as a richness that 'subtly pervades the auditory dimension of experience' (*ibid.*, p. 77). This ambiguity—and indeed, richness—is vividly highlighted within deep listening episodes.

To enact a deep listening episode is precisely to play with the ambiguity at the heart of the musical domain's double dimensionality. The initial phase of deep listening might begin by narrowing one's focal attention to capture the sonic shape and texture of a particular sound feature. This initial gesture inaugurates entry into the inner structural space of the music; it is a focused entry into the piece's temporal dynamics. For example, while listening to a favourite piece, the listener may start by attentively following the contour of a melody, listening to and then eventually 'past' its dynamics as it gradually traces a narrative path through musical time and tonal space. This latter notion serves as 'a designation that corresponds to our perception of music as moving *through* something—for example, from a higher position to a lower one' (Morgan 1980, p. 528). Within deep listening, a melody is experienced as unfolding within a spatialized auditory dimension that the deep listener simultaneously moves to inhabit (hence, the experience's immersive character).

As this listening becomes intensified and further focalized—i.e. the listener listens 'past' the pleasant affective solicitations of the melodic contour and becomes aware of things happening behind or below the melody—a more holistic global sensitivity emerges. Another region of inner space becomes phenomenally accessible: the space of musical *relationships*. The experience as of a melody unfolding within tonal space invites the deep listener to become attuned to perceptual differentiations

co-given. But co-giveness is not equivalent to sameness of experiential intensity. The allocentric information of 'shallow' hearing remains experientially focalized; it stands out in a way not the case in deep listening, as the descriptions below will attempt to make clear.

between other things happening in the piece. In other words, this felt appreciation of tonal space amplifies an appreciation of the piece's *texture*—that is, its *density* as comprised of multiple simultaneous sound events. For, attentively focusing on the dynamics of a melody (or other discrete sound units) does not obscure other musical events happening in the piece. Rather, melody and accompaniment, for example, 'do not simply merge into a single temporal continuum but appear to occupy different spatial locations, thus maintaining both individuality and a clear mutual relationship' (*ibid.*, p. 529). So, within this next phase of deep listening, the melody is perceived to unfold within a different region of auditory space than its rhythmic accompaniment; the latter undergirds the former. Both thus trace parallel but distinct paths through tonal space. And in perceiving this distinction, the listener thus becomes acutely aware of the (auditory) spatial relation between melody and rhythmic accompaniment—the relation itself becomes a positive feature of the listener's awareness—further deepening and refining the listener's perception of both melody and rhythmic accompaniment. Put differently, the phenomenal appreciation of the *relation* simultaneously enriches the appreciation of the *relata*. The experience of the piece is thus qualitatively deepened. What began as an attentive inspection of melody has thus gradually shifted to a more subtle appreciation of the inner architectonics of the piece as a whole—an appreciation of the 'aggregate quality' of the various inter-relationships linking musical events together within tonal space (*ibid.*, p. 529). And this subtle shift in quality of attention signals an experiential fusing of inner and outer musical space, a blending of surroundability and directionality. The piece is now inhabited. The listener is in a position to actively explore different aspects of this nested acoustic environment from an inside-out perspective, as it were.

The animate body plays an important role in enacting this sort of musical spatial fusing. This is because bodily movements such as gently swaying back and forth, bobbing one's head, tapping fingers and toes, and of course dancing—more on this in a moment—modulate our perception of the spatial content of musical experience by modulating our relation to different features of the music, such as metre and melody. Bodily gestures in response to musical events can act as a kind of attentional focusing: the animate body, by interactively engaging with the piece, becomes a vehicle for voluntarily drawing out certain features of the piece, such as rhythmic beats or the progression of a melodic contour, by foregrounding them in our attentional field. This 'drawing out' is an enactive and exploratory gesture in response to felt affordances within the music.⁸ The listener per-

8 There is much empirical evidence indicating that we seem to perceive music as affording various forms of bodily interaction from birth. For discussion of the notion of 'musical affordances', see Krueger (forthcoming a).

ceives the inner space of the piece as a space that can be entered into, experientially, and by doing just this shapes how the experiential content of the piece-as-given becomes phenomenally manifest. We thus hear what the body feels (Philips-Silver and Trainor 2007). And what the body feels are sensorimotor contingencies—possibilities for rhythmic interaction and perceptual exploration that determine the character and content of musical experience.

This idea is reinforced by considering shared episodes of deep listening. Consider, for example, how the shared attentional framework in a live music setting affects the group's mutual perception and appropriation of the music. In particular, consider the role that the crowd's latching onto musical textures and sonic patterns plays in shaping the shared experience of music within a live setting. For example, the simple act of a guitarist casually strumming the first few chords of a popular song—especially at the beginning of a concert—immediately elicits a thunderous roar of approval from the audience. Within a moment, the crowd's attention is galvanized around these textures, snapping into a mode of taut expectancy; the atmosphere is flush with anticipation of the song that will soon follow. Once this familiar refrain begins, the texture (and thus the inner space) of a piece is progressively structured by the introduction of new sound elements: the initial guitar strumming is soon girded by the rhythmic pulsing of the bass accompaniment; keyboards emerge to fill in the sound even more, enriching and deepening the sonic structure; next, the drums enter, stabilizing and accelerating the song's forward momentum; finally, vocals materialize, their aurally-discernible human character lending a sense of qualitative unity to the piece as a whole (in addition to whatever narrative dressing the lyrics offer). This gradually-unfurling sonic world invites shared exploration and appropriation. That this is so becomes clear when we observe how the emergence of each new aspect of these textures elicits new bodily responses from the audience (modulations of head bobbing, swaying, and other whole-body movements, dancing, shouts of encouragement, etc.), as well as a collective refocusing of attention on each emergent sound feature as it comes forward. The musicians perceive and respond to these cues—altering their performance accordingly—which in turn subsequently shapes the audience's further responses and attentional refocusing. Within this organic performer-perceiver interplay, then, a shared attentional framework emerges unique to that time and place. The material mediation of this particular music event—the way that the live music event is embodied in things like the number of listeners, the spatial location of the performance, the musical skills of the performers and audience, and the social values of the attendees—determines both (1) what sort of shared attentional framework emerges in that context, and (2) how musical textures are perceived and

appropriated by multiple perceivers via this attentional framework. If the audience were to suddenly disappear, save for one lone listener, the ambient intensity, tension, and attentive focus to particular musical textures would also disappear – and the phenomenal character of the music-as-given to our lone listener would be dramatically altered. Try as we might, we simply cannot recreate this atmosphere within our solitary listening episodes. The experiential character of deep listening is thus dramatically different when others are involved.

To continue with this example: an integral part of many live music experiences is dancing. Dancing is a robustly embodied response to musical events. Moreover, it is the enactment of an attentional focusing that shapes how and what we hear. Via dancing, the temporal regularities of melodic and rhythmical patterns within the music are physicalized within an array of bodily movements. And the coordination between sonic pattern and bodily movement – an instance of bodily entrainment – is an enactive gesture, a perceptual exploration of the piece's sonic topography. Again, bodily movements modulate the listener's relation to different features of the piece (e.g. metre and melody); dancing experientially foregrounds these features and shapes the way that these features stand out against the background of the piece's other sound features. The temporal predictability and consistency of sonic patterns afford this sort of bodily engagement. Sonic patterns therefore afford an entering into the inner recesses of sonic space, a point of access for losing ourselves, experientially, within the piece via the immersive 'deep listening' that often occurs whilst dancing.

Yet dancing at a live music performance is not simply an instance of the listener being aware of and responding to solitary possibilities for musical interaction. Additionally, part of the content of the dancer's awareness is the dancing-responses of *other* dancers. Their attunement and reaction to sonic patterns shapes the listener's *own* experience of these patterns. Dancing is thus a vehicle of joint attention, a means of enacting a shared attentional framework that shapes the character and content of the piece-as-perceived in that context.⁹ When others' dancing reactions shift, for instance, I feel my body compelled to alter my own movements accordingly. I bob my head in time and sway my body, carried along by the pulse and tempo of the crowd's movement. In this way do I come to inhabit the lived time and structural space of the musical piece with others. The rhythmical and sonic patterns of a piece, and the dancing these patterns afford, forge an interactive phenomenon that synchronizes the joint listeners to one moving mass (Vickhoff and Malmgren 2004, p. 19). And this one moving mass enacts a shared attentional framework unique to that time, place, and performance. However, both the individual as well as collective integrity of the experience simultaneously coexist within that musical experi-

⁹ See Cochrane (2009) for more on joint attention and shared musical experience.

ence. Participants within that shared experience are able to interpret a steady flow of musical features and patterns in individual terms, while the temporal regularities of the sonic invariants discussed coordinate their individual behaviour as well as their attentional foci (Cross 2006, p. 122).

This enactive characterization of musical experiences, both solitary and shared, emphasizes the central role that agency plays in shaping both the character and content of musical experience – including its spatial content. Having offered some phenomenological descriptions of different forms of deep listening, I now want to look at supplementary empirical evidence that seems to support these descriptions.

5 Empirical support

The first line of empirical evidence I want to look at concerns amusia. Amusia is profound tone deafness, an inability to hear music as music. More formally, it is a severe deficiency in processing pitch variation despite normal speech perception and intact sense of rhythm (Ayotte *et al.* 2002; Peretz *et al.* 2002; Sacks 2007). There are different forms of amusia. For the total amusic, however, music is experienced as incoherent noise, an irritating sound structure lacking any sort of aesthetically-compelling character. For example, one amusic described Rachmaninov's second piano concerto as sounding like 'banging and noise' (McDonald and Stewart 2008), whereas another describes the experience of listening to music as akin to a screeching car (Sacks 2007, p. 101).

The conventional explanation of amusia portrays it as an auditory deficit (1) related to deficiencies in fine-grained processing of musical pitch variations, and (2) confined to the musical domain and musical abilities (Ayotte *et al.* 2002). However, some recent studies challenge this characterization. They propose instead that amusia is not a specifically sensory-musical deficit but rather a *spatial* deficit – that is, an inability to represent space (Cupchik *et al.* 2001; Douglas and Bilkey 2007; Särkämö *et al.* 2009). For instance, amusics were found to perform significantly worse than non-amusic controls on mental rotation tasks (Douglas and Bilkey 2007). Cupchik *et al.* (2001) found a correlation between performance on a mental rotation task involving three dimensional figures and the ability of the listener to perceive inverse and retrograde musical permutation (i.e. when a musical piece had been played backwards). Whether or not amusia stems from a spatial deficit is a matter of some debate (see, e.g. Tillmann *et al.* 2010). However, if something like this is the case, it lends insight into the amusic's musical phenomenology – or rather, lack thereof. For, it seems that amusics are unable to perceive music as offering up the spatially-inviting auditory profile that normal listeners perceive. They might perceive the outer or *locational* spatial profile of music specifying its egocentric location (that 'banging and noise' is coming from *over there*). But they

are unable to perceive a piece's inner or *structural* spatial profile – that is, the spatial quality that specifies that sound event's uniquely musical character. And without the ability to enact the spatial fusing involved in deep listening, music thus remains an alien and impenetrable entity.

This view receives support from another study. As discussed above, the animate body plays a central role in the spatial fusing characteristic of deep listening. Again, bodily movements such as swaying back and forth, nodding our heads, tapping fingers and toes, or the more energetic whole-body dynamics of dancing, modulate our perception of the spatial dimensions of musical experience. In particular, bodily synchronization with rhythmic patterns and tempo – actions that, as we'll see in a moment, we seem born ready and able to enact – open up the inner space of a piece. This bodily engagement with music is both a response to and an affirmation of music's spatially-structured exploratory profile. Amusiacs, however, have a marked difficulty in synchronizing bodily movements with music – despite a normal ability to synchronize with sequences of non-musical sounds (Dalla Bella and Peretz 2003). Another more recent study affirmed this result, indicating that the deficit in processing rhythmic patterns was related not to the complexity of the patterns themselves – the subjects were able to synchronize with monotonic sounds such as a steady drum beat, for example – but rather the pitch-variations of the music (Foxton *et al.* 2006). Again, lacking the ability to perceive and respond to the inner structural space of the musical piece, the amusiacs were accordingly unable to enact a robust sensorimotor response to the music – which in turn affected both the experiential character of the music-as-perceived (i.e. as having a disagreeable sonic character) as well as the music-as-experiential content (i.e. as an impenetrable object lacking a spatially-inviting exploratory profile).

Daniel Vastfjall (2003) found that both the experienced presence of music (i.e. sound immersion, or the feeling of involvement with a piece) as well as experienced emotions in response to music varies as a function of its perceived spatiality. In other words, the spatial content of musical experience, in contrast to other acoustic parameters (e.g. pitch, timbre, loudness, etc.), is arguably what triggers the profound immersive and emotional responses characteristic of deep listening. In Vastfjall's study, participants were seated and asked to close their eyes and listen in a concentrated (i.e. 'deep') way – they were told to 'let themselves into the music' – focusing in particular on the intensity of their emotional reactions to the pieces (Vastfjall 2003, p. 184). Predictably, participants in the 'mono condition' (i.e. listening to music via one-channel) did not respond to emotion induction to the same extent as participants in either the stereo (i.e. two-channel, or speakers on either side) or six loudspeakers conditions (*ibid.*, p. 185). In the latter conditions, the music was experienced as more

immersive and thus more emotionally compelling. It appears, then, that the 'subjective sense of presence and emotional reactions to the music are highly interrelated' – affirmed by the fact that '[p]articipants who experienced a strong feeling of presence and a sense of being in the sound field also reported stronger emotional reactions' (*ibid.*, p. 186). When the spatial content of musical experience is absent or diminished – such as with amusia, or perhaps in more common inattentive or shallow modes of listening – the immersive and emotional character is also compromised. Likewise, Don Ihde quotes a philosopher friend who recalls first becoming aware of his increasing deafness when he began to lose interest in music, which gradually became 'distant... objectlike... over there apart from me' (Ihde 2007, p. 78). In this case, music appears to have lost its spatial character; what was once a dynamic, spatially-structured soundworld was reduced to an inert acoustical object. Deep listening was no longer possible, only observational hearing.

With practice and experience, one can presumably cultivate and refine the attentional and sensorimotor skills needed for deep listening. However, multiple streams of empirical evidence from neonate music therapy suggest that we are potentially deep listeners from birth. This is not the place for a comprehensive review of the literature (see, e.g. Standley 2001; for discussion, see Krueger forthcoming a,b). But we can note a few salient points. Generally speaking, music therapy consists of a cluster of various music-related practices and techniques designed to give patients of all ages the opportunity to explore and communicate emotions (Bunt and Pavlicevic 2001, p. 181). Traditionally geared toward children and adults with various disabilities or mental health problems, the past few decades have seen a rising interest in music's therapeutic effect on neonates. Specifically, neonate music therapy has arisen in response to what Tia DeNora terms the 'paradox of the NICU' (DeNora 2000). This is the idea that the hostile soundworld of the NICU – comprised of, for example, the auditory byproduct of medical technologies (e.g. respirators, bottles clanking on incubators, noisy beeps of heart monitors and other machinery amplifying aspects of the infant's disorganized state, etc.), the sound of other infants crying, the continual commotion of people moving in and out of the area – might actually be disrupting the infant's basic life-processes, in turn affecting sleep regularization and state lability (*ibid.*, p. 80; see also Haslbeck 2004; Kaminski and Hall 1996, p. 46). However, a significant amount of research seems to indicate that music can be a valuable resource for enhancing the neonate's physiological and emotional well-being, serving as an occluding corrective to this unfriendly sound environment (Standley 2001, p. 213; DeNora 2000, p. 81).

For the purposes of this discussion, I am most interested in how neonates and infants enact musical experience – that is, how infants seem to

perceive music as presenting a spatially-structured exploratory profile inviting bodily entrainment. From the start, infants are surprisingly skilled listeners, seemingly attuned to the rhythmic, emotional, and communicative opportunities that musical engagement offers.¹⁰ Like adults, they appear to appreciate and respond to music as an experientially salient feature of their perceptual environment. For example, both term and pre-term infants attend more fixedly to music than they do to other ambient noises, suggesting a preference for the sonic coherence and organizational structure of music in contrast to contingent environmental noise (Butterfield and Siperstein 1972; Standley 2001). Infant activity tends to decrease in response to auditory stimuli generally. But the most significant decreases are caused by music, further suggesting that music is a preferred auditory stimulus (Kagan and Lewis 1965). Other studies have found that infants are surprisingly discriminating listeners. Not only do infants tune in to overarching musical patterns, preferring consonant over dissonant intervals (Trainor and Heinmeiller 1998, p. 83; Zentner and Kagan 1998). Additionally, they are able to pick out and attend to fine-grained auditory properties of music such as pitch, melody, tempo, and musical phrase structure (Schellenberg and Trehub 1996; Trehub *et al.* 1999; Trehub and Schellenberg 1995; Trehub and Trainor 1993). For example, three- to six-month-olds can vocalize a matched pitch to sung tones (Wendrich 1981) as well as learn to turn toward a loudspeaker whenever they perceive a change in background melody (Trehub *et al.* 1987). Two-month-olds can remember short melodies and discriminate it from other melodies (Plantinga and Trainor 2009). Infants, it would seem, are therefore capable of hearing and responding to the particular sound features that carry a piece's expressive content. Even the very young possess the perceptual skills needed to find music perceptually captivating because of its emotional expressivity (Nawrot 2003). Additionally, they possess the listening skills needed to actively *explore* music, to *enact* musical experience, by selectively attending to and bodily engaging with aspects of its sonic topography.

To further see how this is so, consider that, beyond merely exhibiting the perceptual skills needed to make musical discriminations, infants also seem to experience music as affording communicative possibilities. This is indicated by their rhythmic bodily entrainment responses to music. Haslbeck (2004), for example, found that, over the course of several music therapy sessions, the pre-term neonates in her study gradually became active participants within the sessions, intentionally seeking interpersonal contact via the music, which consisted of slowly-sung melodies supplemented with a hand resting gently on the infant's chest or back. This inter-

¹⁰ Again, considerations of space prohibit an adequate defence of this thesis here. See instead Krueger (forthcoming b).

personal contact emerged via bodily entrainment: the infants enacted whole-body ‘rhythmic dialogues’ with the music (*ibid.*, p. 9). These dialogues were established via a coordinated rhythmic alteration between the sung lullaby and the infants’ bodily responses. For instance, both sucking/swallowing and regularized patterns of respiration were observed to mimic the rhythmic alterations of the sung melody (e.g. sucking at the end of melodic phrases – *ibid.*, p. 9). More tellingly, the infants gradually initiated eye contact, summoned an increasingly attentive and engaged posture (while reducing fidgeting and grimacing), and exhibited increased mouth movements (playing with the tongue, mouthing the vowels being sung, such as ‘o’ and ‘u’) and vocalizations during the sessions. Other movements included opening and closing of hands, wrinkled brows, and eyes opening and closing in sync with the rising of falling of the sung melody (*ibid.*, p. 11). This opportunity for social contact within music therapeutic contexts is crucial – and, seemingly, something that the infants in Haslbeck’s sessions actively sought out – given the affective isolation of the pre-term infant’s life inside the incubator. Additionally, it suggests that infants are enactively attuned to the spatial characteristics of music since, without this attunement – as the amusia research demonstrated – this sort of rhythmic entrainment cannot occur.¹¹

What the evidence discussed above indicates, I suggest, is that particular structural features of music – again, textural qualities and regularities of melodic and rhythmic patterns – are actively perceived and exploited by infants in episodes of proto-deep listening.¹² This is further confirmed by research indicating that movement influences auditory encoding of rhythm patterns in both infants and adults. How we move shapes both what we hear and how we hear it. In a series of experiments, Jessica Philips-Silver and Laurel Trainor trained 7-month-old infants by listening to an ambiguous two-minute rhythmic pattern (i.e. a pattern lacking accented beats). Half of the infants were bounced on every second beat and half were bounced on every third beat. As a result, the infants expressed a more prolonged interest in the auditory test stimulus with the metrical form – every second beat accented (the duple form) in one stimulus, and every third beat (the triple form) in the other – that matched the metrical form of their training bouncing (Philips-Silver and Trainor 2007, p. 1430). This was also the case when blindfolded. A further experiment

- 11 Another study found that preverbal infants spontaneously display tempo-sensitive rhythmic motions of their body with music – and that this bodily engagement is a source of positive affect – but that they don’t exhibit this behaviour in response to speech or other arrhythmic ambient noise (Zentner and Eerola 2010). There are also indications that some animals (e.g. parrots) are capable of similar coordination (Patel *et al.* 2009; Schachner *et al.* 2009). Perhaps they, too, are capable of a kind of proto-deep listening.
- 12 Gentle music with ‘thin’ textures (e.g. sung lullabies, new age music, etc.) and no abrupt modulations of volume or tempo reduces alerting responses in infants, offering up a stable, inviting, and predictable soundworld for the infant to explore (Standley 2001; 2002).

showed that personal bodily movement was necessary to establish this metrical preference. Watching the experimenter bounce during the ambiguous rhythm training failed to establish a preference for either of the auditory stimulus versions (*ibid.*, p. 1430).¹³ A similar set of experiments was later done with adults (*ibid.*). Unlike the infants, of course, the adults could engage in their own ‘bounce training’. But like the infants, the adults’ synchronized movements of their body determined how they heard an ambiguous musical rhythm (*ibid.*, p. 543). Once again, they had to personally bounce their own bodies, and not watch a video of another doing it, in order for their experience of the ambiguous rhythm to covary relative to their particular bounce training (e.g. bouncing on every second or on every third beat). But their sensorimotor training determined how they enacted the content of their experience of the ambiguous rhythm. Ample empirical evidence therefore suggests that even infants possess rudimentary (i.e. practical or pre-theoretical) sensorimotor understanding of how modulations of bodily movement and attentional focusing affect sensory change. They are capable of enacting rich musical experiences from the start. Moreover, music is perceived, again from the start, in terms of its spatial character. For the normal listener, music manifests experientially as harbouring non-derived spatial content.

6 An objection

I now want to briefly consider a natural objection to the enactive view defended above. We can term this the ‘immobile listener objection’. This is the objection that listeners with various sorts of extreme sensorimotor deficits (e.g. quadriplegics, individuals with Locked-in Syndrome, etc.) lack the ability to enact a robust sensorimotor engagement with music. Yet they nevertheless clearly perceive music as music – and surely, moreover, are capable of being moved deeply by it, experiencing it in an immersive and emotionally resonant way. Thus, music listening cannot depend essentially upon exploratory sensorimotor skills and actions the way this chapter has argued that it does.

However, this objection misses the mark for a couple of reasons. First, many sufferers of spinal cord injuries or paralysis had extensive periods of perceiving prior to suffering their injury (e.g. as the result of a fall or car accident) or the onset of, for example, Multiple Sclerosis in young adulthood or Locked-in Syndrome later in life. So, they clearly retain a practical understanding of how movement and attentional focusing modulates sensory change – even once their movement is inhibited. Moreover, they retain a range of practical skills needed to enact experience of different sorts, such as the ability to move their eyes, head, and (with the assistance

¹³ For samples of the experimental sound stimuli, see the following link: <http://www.sciencemag.org/cgi/content/full/308/5727/1430>

of technology such as a wheelchair) their entire bodies in relation to their environment. Despite their physical limitations, quadriplegics lead active exploratory lives. They are continually 'engaged in the task of orienting themselves in relation to the world around them and to gravity' (Noë 2004; see also Cole 2004). Thus, they remain active perceivers even if the range of their active perceiving is somewhat restricted. And in the case of music perception, those in wheelchairs are entirely capable of enacting rhythmic synchronization with music – which, as we've seen, is a crucial enactive gesture for opening up the spatial character of the auditory event *qua* music, a process which affords deeper, more focalized listening. Not only can they nod their heads or blink in time with music, or attentionally follow the contour of a melody as it moves through tonal space. Additionally, wheelchairs can be summoned to perform all sorts of skilled, active engagements with music – swaying back and forth, twirling in circles, tilting from side to side, etc. – as an internet video search will quickly reveal. These movements allow the wheelchair-bound listener to explore how the dynamics of embodied engagement alter the character and content of musical experience. Wheelchair-bound perceivers are thus far from immobile listeners. To the contrary, they remain capable of enacting rich spatially-structured musical content. They retain the skills and practical understanding needed to respond to the unique exploratory profile music offers; they have the skills to access music in a sensitive or 'deep' manner. It is rather those with a *spatial* deficit (e.g. amusiaks) who can no longer perceive and respond to this exploratory profile and thus who have lost this ability. In the latter cases, both the character and content of musical experience has been dramatically altered, as the perceptual reports of amusiaks would seem to indicate.

7 Concluding thoughts

I have argued that listening to sounds is an active sensorimotor exploration of our world – including spatial and locational aspects of that world and things in it. Sounds routinely furnish spatial information about our world, and we use our auditory experiences to explore and skillfully respond to this information. This is particularly evident in the case of music perception. Specifically, I have argued that we enact our musical experience – that is, we summon a range of bodily skills to secure experiential access to music, and, in particular, its spatially-complex character. For it is here that the source of music's experiential richness lies. This complex spatial character – as well as the way that this complex spatial character is articulated in and through music's temporal dynamics – is what makes music such a uniquely compelling phenomenon. I have tried to bring this out with a phenomenological characterization of what I termed 'deep listening', focusing in particular on how bodily skills and active per-

ceptual exploration play a central role in enacting musical content. The account developed above is, of course, merely a sketch. But for enactive accounts of perception to develop, they must extend the discussion beyond the well-tread terrain of vision and touch and move into the domain of other sensory modalities. Considering the nature of music perception, I suggest, is a particularly fertile way to do just this.

References

- Ayotte, J., I. Peretz and K. Hyde (2002), 'Congenital Amusia: A Group Study of Adults Afflicted with a Music-Specific Disorder', *Brain*, **125**: 238–51.
- Becker, J. (2004), *Deep Listeners: Music, Emotion, and Trancing*, Bloomington (IA), Indiana University Press.
- Bunt, L. and M. Pavlicevic (2001), 'Music and Emotion: Perspectives from Music Therapy' in P.N. Juslin and J.A. Sloboda, Eds., *Music and Emotion: Theory and Research*, Oxford, Oxford University Press.
- Butterfield, E. and G. Siperstein (1972), 'Influence of Contingent Auditory Stimuli Upon Nonnutritional Suckle' in J. Bosma, Ed., *Oral Sensation and Perception: The Mouth of the Infant*, Springfield (IL), Charles C. Thomas: 313–33.
- Casati, R. and J. Dokic (2009), 'Some Varieties of Spatial Hearing' in M. Nudds and C. O'Callaghan, Eds., *Sounds and Perception: New Philosophical Essays*, Oxford, Oxford University Press: 97–110.
- Cochrane, T. (2009), 'Joint Attention to Music', *British Journal of Aesthetics*, **49** (1): 59–73.
- Cole, J. (2004), *Still Lives: Narratives of Spinal Cord Injury*, Cambridge (MA), MIT Press.
- Cross, I. (2006), 'Music and Social Being', *Musicology Australia*, **XXVIII**: 114–26.
- Dalla Bella, S. and I. Peretz (2003), 'Congenital Amusia Interferes with the Ability to Synchronize with Music', *Annals New York Academy of Sciences*, **999**: 166–9.
- DeNora, T. (2000), *Music in Everyday Life*, Cambridge, Cambridge University Press.
- Douglas, K.M. and D.K. Bilkey (2007), 'Amusia is Associated with Deficits in Spatial Processing', *Nature Neuroscience*, **10** (7): 915–21.
- Foxton, J.M., R.K. Nandy and T.D. Griffiths (2006), 'Rhythm Deficits in "Tone Deafness"', *Brain and Cognition*, **62** (1): 24–9.
- Gangopadhyay, N. and J. Kiverstein (2009), 'Enactivism and the Unity of Perception and Action', *Topoi*, **28** (1): 63–73.
- Hamilton, A. (2009), 'The Sound of Music' in M. Nudds and C. O'Callaghan, Eds., *Sounds and Perception: New Philosophical Essays*, Oxford, Oxford University Press: 146–82.
- Haslbeck, F. (2004), 'Music Therapy with Preterm Infants – Theoretical Approach and First Practical Experience', *Music Therapy Today (online)*, **5** (1): 1–15.
- Hurley, S. (2001), 'Perception and Action: Alternative Views', *Synthese*, **129** (1): 3–40.
- Hyde, P.S. and E.I. Knudsen (2002), 'The Optic Tectum Controls Visually Guided Adaptive Plasticity in the Owl's Auditory Space Map', *Nature*, **415** (6867): 73–6.
- Ihde, D. (2007), *Listening and Voice: Phenomenologies of Sound*, Albany (NY), SUNY Press.
- Kagan, J. and M. Lewis (1965), 'Studies of Attention in the Human Infant', *Journal of Developmental Psychology*, **11** (2): 95–127.
- Kaminski, J. and W. Hall (1996), 'The Effect of Soothing Music on Neonatal Behavioral States in the Hospital Newborn Nursery', *Neonatal Network*, **16**: 45–54.
- Kierkegaard, S. (1959), *Either/Or* (Vol. 1), D.F. Swenson, Trans., Garden City (NY), Double Day.
- Krueger, J. (forthcoming a), 'Doing Things with Music', *Phenomenology and the Cognitive Sciences*.
- Krueger, J. (forthcoming b), 'Empathy, Enaction, and Shared Musical Experience' in T. Cochrane, B. Fantini and K.R. Scherer, Eds., *The Emotional Power of Music: Multidisciplinary Perspectives on Musical Expression, Arousal and Social Control*, Oxford, Oxford University Press.

- McDonald, C. and L. Stewart (2008), 'Uses and Functions of Music in Congenital Amusia', *Music Perception*, **25** (4): 345-55.
- Merleau-Ponty, M. (2004), *The World of Perception*, New York, Routledge.
- Morgan, R.P. (1980), 'Musical Time/Musical Space', *Critical Inquiry*, **6** (3): 527-38.
- Noë, A. (2004), *Action in Perception*, Cambridge (MA), MIT Press.
- Noë, A. (2009), 'Conscious Reference', *The Philosophical Quarterly*, **59** (236): 470-82.
- Nonken, M. (2008), 'What Do Musical Chairs Afford? On Clarke's Ways of Listening and Sacks's Musicophilia', *Ecological Psychology*, **20** (4): 283-95.
- Nudds, M. (2001), 'Experiencing the Production of Sounds', *European Journal of Philosophy*, **9** (2): 210-29.
- O'Callaghan, C. (2009), 'Auditory Perception' in E.N. Zalta, Ed., *The Stanford Encyclopedia of Philosophy*, [Online], <http://plato.stanford.edu/archives/sum2009/entries/perception-auditory/>
- O'Callaghan, C. (2010), 'Perceiving the Locations of Sounds', *Review of Philosophy and Psychology*, **1** (1): 123-40.
- O'Regan, J.K. and A. Noë (2001), 'A Sensorimotor Account of Vision and Visual Consciousness', *Behavioral and Brain Sciences*, **24** (5): 939-73.
- O'Shaughnessy, B. (1984), 'Seeing the Light', *Proceedings of the Aristotelian Society*, New Series, **85**: 193-218.
- Patel, A.D., J.R. Iversen, M.R. Bregman and I. Schulz (2009), 'Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal', *Current Biology*, **19** (10): 827-30.
- Peretz, I., J. Ayotte, R.J. Zatorre, J. Mehler, P. Ahad, V.B. Penhune and B. Jutras (2002), 'Congenital Amusia: A Disorder of Fine-Grained Pitch Discrimination', *Neuron*, **33** (2): 185-91.
- Plantinga, J. and L.J. Trainor (2009), 'Melody Recognition by Two-Month-Old Infants', *The Journal of the Acoustical Society of America*, **125** (2): 58-62.
- Philips-Silver, J. and L.J. Trainor (2007), 'Hearing What the Body Feels: Auditory Encoding of Rhythmic Movement', *Cognition*, **105** (3): 533-46.
- Putnam, H. (1999), *The Threefold Cord: Mind, Body, and World*, Columbia (NY), Columbia University Press.
- Keybrouck, M. (2005), 'A Biosemiotic and Ecological Approach to Music Cognition: Event Perception Between Auditory Listening and Cognitive Economy', *Axiomathes*, **15**: 229-66.
- Sacks, O. (2007), *Musicophilia: Tales of Music and the Brain*, London, Picador.
- Särkämö, T., M. Tervaniemi, S. Soinila, T. Autti, H.M. Silvennoinen, M. Laine and M. Hietanen (2009), 'Cognitive Deficits Associated with Acquired Amusia After Stroke: A Neuropsychological Follow-Up Study', *Neuropsychologia*, **47** (12): 2642-51.
- Schachner, A., T.F. Brady, I.M. Pepperberg and M.D. Hauser (2009), 'Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species', *Current Biology*, **19** (10): 831-6.
- Schellenberg, E.G. and S.E. Trehub (1996), 'Natural Musical Intervals', *Psychological Science*, **7** (5): 272-7.
- Schopenhauer, A. (1966), *World as Will and Representation*, E. Payne, Trans., (Vol. 1), Mineola (NY), Dover Publications.
- Sloboda, J.A. (2005), *Exploring the Musical Mind: Cognition, Emotion, Ability, Function*, Oxford, Oxford University Press.
- Small, C. (1998), *Musicking*, Middletown (CT), Wesleyan University Press.
- Smith, F.J. (1979), *The Experiencing of Musical Sound: Prelude to a Phenomenology of Music*, New York, Gordon and Breach.
- Standley, J. (2001), 'Music Therapy for the Neonate', *Newborn Infant Nursing Reviews*, **1** (4): 211-6.
- Standley, J.M. (2002), 'A Meta-Analysis of the Efficacy of Music Therapy for Premature Infants', *Journal of Pediatric Nursing*, **17** (2): 107-13.
- Thompson, E. (2005), 'Sensorimotor Subjectivity and the Enactive Approach to Experience', *Phenomenology and the Cognitive Sciences*, **4** (4): 407-27.

- Tillmann, B., P. Jolicœur, M. Ishihara, N. Gosselin, O. Bertrand, Y. Rossetti and I. Peretz (2010), 'The Amusic Brain: Lost in Music, but Not in Space', *PLoS ONE*, **5** (4): e10173.
- Trainor, L.J. and B.M. Heinmiller (1998), 'The Development of Evaluative Responses to Music: Infants Prefer to Listen to Consonance Over Dissonance', *Infant Behavior and Development*, **21** (1): 77-88.
- Trehub, S.E. and E.G. Schellenberg (1995), 'Music: Its Relevance to Infants', *Annals of Child Development*, **11**: 1-24.
- Trehub, S.E., E.G. Schellenberg and S. Kamenetsky (1999), 'Infants' and Adults' Perception of Scale Structure', *Journal of Experimental Psychology*, **25** (4): 965.
- Trehub, S.E. and L. Trainor (1993), 'Listening Strategies in Infancy: The Roots of Music and Language Development' in S. McAdams and E. Bigand, Eds., *Thinking in Sound: The Cognitive Psychology of Human Audition*, Oxford, Oxford University Press.
- Turner, P., S. Turner and I. McGregor (2007), 'Listening, Corporeality and Presence', presented at the *Presence 2007: 10th Annual Workshop on Presence*, Barcelona, Spain, [Online], <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.92.7329>
- Vastfjall, D. (2003), 'The Subjective Sense of Presence, Emotion Recognition, and Experienced Emotions in Auditory Virtual Environments', *CyberPsychology & Behavior*, **6** (2): 181-8.
- Vickhoff, B. and H. Malmgren (2004), 'Why Does Music Move us?', *Philosophical Communications, Web Series*, **34**, Dep. of Philosophy, Goteborg University, Sweden.
- Zahorik, P. and F.L. Wightman (2001), 'Loudness Constancy with Varying Sound Source Distance', *Nature Neuroscience*, **4** (1): 78-83.
- Zentner, M.R. and T. Eerola (2010), 'Rhythmic Engagement with Music in Infancy', *Proceedings of the National Academy of Sciences of the United States of America*, [Online], <http://www.pnas.org/content/early/2010/03/08/1000121107.full.pdf+html?frame=header>
- Zentner, M.R. and J. Kagan (1998), 'Infants' Perception of Consonance and Dissonance in Music', *Infant Behavior and Development*, **21** (3): 483-92.