

## Sonic Pictures

{Winning essay of the American Society for Aesthetics' inaugural Peter Kivy Prize.}

Jason Leddington

Department of Philosophy, Bucknell University  
Centre for Philosophical Psychology, University of Antwerp

In *Antithetical Arts*, Peter Kivy explains how absolute music—"an art of purely abstract but perhaps expressive sound"—poses a distinctive problem for the philosophy of art:

All of the *other* fine arts are, for the most part, arts *with* literary or representational content. And that content plays a major role in accounting for what it is in these arts that gives us such deep and abiding satisfaction. But absolute music does not possess such content. So it is a puzzle as to what it is in or about absolute music gives what *appears*, at least, to be the same *kind* of deep satisfaction that the other arts, the arts with content, give. That, in brief, is the "problem" of absolute music. (2011, 119)

One obvious solution to the problem, "popular both in philosophical and music-theoretical circles," is to contest appearances and "to deny that absolute music does indeed want for literary [or representational] content" (2011, 119). Following arch-formalist Eduard Hanslick, Kivy thinks such views are hopeless (2011, chaps. 3–7). Instead, he insists that a fully satisfactory account of the artistic and aesthetic satisfactions of "music alone" *must*—and *can*—be given without appeal to any sort of literary or representational content (1990; 2011, chaps. 8–11). In this respect, literature and absolute music are indeed "antithetical arts."<sup>1</sup> However, the same cannot be said of music in general. Despite his formalism, Kivy is deeply sensitive to the representational potential of music, which he explores at length in *Sound and Semblance* (1984) and *Osmin's Rage* (1988). Focusing primarily on program music and opera, Kivy argues that there are many cases in which full appreciation of a work or performance requires attention to the music's representational features. But more controversially, against Roger Scruton (1976),

---

<sup>1</sup> However, Kivy also rejects Hanslick's "extreme formalism," according to which "[t]he content of music is tonally moving forms, empty of any other content, emotive content in particular" (2011, 64; quoting Hanslick 1986, 28–9). According to Kivy, Hanslick's mistake was to fail to grasp the possibility of an "enhanced formalism," "a formalism that recognizes emotive properties of music as perceptual, phenomenological properties, not semantic or representational ones" (2011, 60, 74; but cf. 64–5).

## Sonic Pictures

Jenefer Robinson (1987), and Stephen Davies (1993), Kivy claims that, in rare cases, music, like painting, can actually provide *pictorial* experience. This is my topic here.

However, my concern is not the *possibility* of musical picturing; I believe that Kivy's arguments—including his replies to Scruton, Robinson, and Davies—conclusively establish it (see Kivy 2012). Instead, my goal is to build on Kivy's work to develop a conception of musical pictures—or, more generally, “sonic pictures” (2012, 152)—that is sensitive to recent developments in philosophy of perception. As it turns out, Kivy's approach to sonic pictures embeds a commitment to a metaphysics of sounds and hearing that significantly restricts the scope of what can be sonically pictured, and recent work in philosophy of perception suggests that we have good reason to question this commitment and the restriction it entails. In its place, I'll recommend a view of sounds and hearing that yields a considerably more powerful conception of sonic picturing.

The article is in four parts. Part I presents Kivy's view of sonic pictures, exposes its metaphysical commitments, and identifies two ways to resist them. Part II introduces five contemporary views that reject the metaphysical basis of Kivy's restriction on sonic picturing and presents an argument in favor of my preferred view. Part III then introduces and responds to a musically motivated objection originally developed by Roger Scruton, which insists that reflecting on musical listening should lead us to prefer something like Kivy's view of sounds. Finally, Part IV explores the view of sonic picturing entailed by the metaphysics of sounds and hearing that I recommend. Among other things, I'll argue that audio tracks on film and two contemporary musical genres—beatboxing and cover songs—are best understood in terms of sonic picturing.

### I. Kivy on Musical Pictures and the Nature of Sound

In exploring the representational capacities of music, Kivy distinguishes “pictorial” from “structural” representations (2002, 183; cf. 1984, chap. 2). In the latter, a “structural element in the music corresponds with something” extra-musical, such as a feature of the accompanying text, “that the structure, so to say, analogizes” (2002, 190). For example, according to Kivy, the resolution to D major at the end of Mozart's *Marriage of Figaro* structurally represents “the resolving of differences among the couples” (2002,

190). But while the informed listener may appreciate this representational function, there is, Kivy notes, “no question” of simply *hearing* conjugal harmony in the music:

There is nothing to ‘hear in’ the music, because the music, in these instances, does not represent sound, or sound events, but abstract concepts and things seen but not heard. One hears the musical structure, understands the text, and perceives, cognizes, the structural analogy. (2002, 190)

By contrast, following Richard Wollheim (1980; 2003), Kivy treats a representation as pictorial just in case it allows us to *perceive* the object of representation *in* the representation itself (2002, 183). This notion of picturing is sense-modality-independent and extends, in principle, beyond the visual. Thus, Kivy writes:

Now visual pictorial representations represent what is *seen*: we *see* the woman *in* the *Mona Lisa*. Likewise, pictorial representations in music, if indeed there are any, represent what is *heard*: we *hear* in the music whatever it pictorially represents. So it seems clear that pictorial representations, if any, in music, must be representations of *sounds*. This does not mean music cannot represent other things besides sounds, or paintings things other than sights. But they can’t represent them pictorially. (2002, 184; cf. 1984, chap. 2)

Call the view that “pictorial representations, if any, in music, must be representations of sounds” *Kivy’s Restriction*. Kivy’s Restriction is grounded in a view of sounds and hearing that Kivy infrequently articulates and (to my knowledge) never defends. It is view with rich philosophical pedigree, and it is shared by Kivy’s main opponent on issues of musical representation, Roger Scruton.<sup>2</sup> I return to it in a moment; first, though, a bit more about Kivy’s view of musical pictures.

Kivy divides pictorial representations into two basic categories: “aided” and “unaided” (2002, 184). An aided pictorial representation is a picture accompanied by words that tell us it is a picture and without which we would have probably failed to recognize it as such. By contrast, we experience unaided pictorial representations as pictures—and can perceive their subjects in them—without having to be told that they are pictures. Kivy writes:

---

<sup>2</sup> On issues of musical representation, Kivy and Scruton disagree across the board. Scruton denies that pure music has *any* representational powers (1976). Against Scruton, Kivy argues that, while musical picturing is uncommon, it can occur, and structural representations in music are commonplace (1984; 1988; 2002; 2012).

## Sonic Pictures

Anyone can see the woman's face in the *Mona Lisa* without being told that it is a portrait of a lady. But there is a beautiful painting by the British artist J. M. W. Turner (1775–1851) in which we can see a sunset in the painting only if we know the title: *Sunset over Lake*. . . . Without the words there would be no 'seeing in': only the impression of a non-representational color composition. (2002, 184)

In the case of visual art, the aided pictorial representation is the exception; but, according to Kivy, in music, it's the rule. He claims that examples of aided musical pictures "abound," but "it is very hard to come up with any real, incontestable examples" of the unaided type (2002, 185, 184). Even the most obvious cases of musical picturing in the Western classical tradition—namely, representations of bird calls in compositions such as Beethoven's Sixth Symphony—are generally flagged as such by titles and other accompanying text. In fact, Kivy claims that "it is probably best to give up the point and admit that unaided pictorial representation in music is, if possible at all, too rare a phenomenon to be counted as belonging to music's repertoire of aesthetic possibilities" (2002, 185). I think it is understandable that Kivy would arrive at this conclusion given his focus on classical music. However, in Part IV, I'll argue that unaided musical picturing is both more common *and* more aesthetically significant than Kivy admits.

While Kivy seems interested only in musical pictures, he clearly intends his account to generalize to "sonic pictures" as such (see esp. 2012, sec. 8.3). Thus, Kivy's Restriction, in full generality, is the view that sonic pictures (musical or otherwise) must be pictures of sounds. Now, I have said that this embeds non-trivial metaphysical commitments. How so? Well, in general, pictures represent what they do, in part, by presenting perceptual appearances shared by their objects. What makes a Rembrandt self-portrait a picture of Rembrandt is, in part, that the portrait shares (enough of) Rembrandt's visual appearance. This, above all, is what enables us to see Rembrandt in the picture. Any satisfactory theory of depiction must offer an account of exactly how—and in what sense—pictures capture the perceptual appearances of the objects they depict.<sup>3</sup> But the point here is that *only* things that perceptually appear—that is, things that *have* perceptual appearances—can be pictured at all (cf. Hopkins 1998, 28–30). At the same time, it seems that *whatever* perceptually appears is something that can in principle be pictured. In sum, then: for sense modality M, an object O might be perceived in an M-

---

<sup>3</sup> See Hopkins (1998) and Kulvicki (2006) for two very different approaches to this task.

## Sonic Pictures

type pictorial representation if and only if O has M-type perceptual appearances. For example, many ordinary objects and events (cabbages, cats, clouds) have visual appearances; thus, they might be seen in visual pictorial representations—that is, they might be visually pictured. Similarly, anything with auditory appearances will be a candidate for sonic picturing—for being heard in sonic pictures. Consequently, Kivy's Restriction—that sonic pictures *must be* pictures of sounds—holds if only if auditory appearances belong *only* to sounds.

The view that auditory appearances belong only to sounds is a piece of our collective empiricist inheritance. As such, it easily masquerades as philosophical common sense. Consider the following exchange, which occurs early in the first of Berkeley's *Three Dialogues*:

PHILONOUS. This point then is agreed between us, that sensible things are those only which are immediately perceived by sense. You will farther inform me, whether we immediately perceive by sight any thing beside light, and colours, and figures: or by hearing, any thing but sounds: by the palate, any thing beside tastes: by the smell, beside odours: or by the touch, more than tangible qualities.

HYLAS. We do not. ([1713] 1992, 138)

In short, we immediately or directly perceive only sensible qualities, which, Philonous goes on to argue, exist only insofar as they are perceived. Such a view receives little support from contemporary philosophers. It is widely agreed that what we immediately perceive are not mind-dependent qualities, but mind-independent objects. In particular, what we immediately see and touch are supposed to be ordinary objects such as horses and tomatoes. This is not to deny that we see colors and shapes; it is to deny that we see the horse by or in virtue of seeing its color or shape. Our visual experience of the horse is not “mediated” by the experience of its sensible properties. Similar considerations hold for touch.

Yet the priority accorded to ordinary objects in visual and tactile perception is typically not extended to the other sense-modalities. In the case of hearing, this manifests itself in two ways. First, philosophers typically follow Berkeley in taking the only direct or immediate objects of hearing to be sounds (for instance, O'Callaghan 2007, 13; 2008, 318; 2009a, 609). On this view, when you witness a musical performance, you directly see the performers, but you directly hear only the sounds they make. Second, even if most

philosophers reject the Berkeleyan view that sounds are mind-dependent qualities, there is still a tendency to think of sounds as independent of their material causes, as somehow hovering above or alongside the world of everyday material objects and events (Strawson 1959, chap. 2; Scruton 1997, chap. 1; 2009; Nudds 2001; 2010; 2014; Martin 2012). Taking these two ideas together suggests that the only items that *can* show up in auditory consciousness—and so, bear auditory appearances—are sounds (Nudds 2014). Thus, in listening to the musical performance, you may (indirectly) hear the performers, their instruments, and their playing, but the only things that genuinely populate your auditory consciousness are the sounds that they make. In this case, no musician, instrument, or act of playing has ever *auditorily appeared* to anyone.

As we'll see in a moment, the trend in the philosophy of perception is to reject such a conservative view of auditory appearances, and this in turn opens the door to a more powerful conception of auditory picturing (among other things). First, however, a key point: if sounds are the sole bearers of auditory appearances, then sounds must be *individuals* rather than *sensible properties*.<sup>4</sup> Here's why. In general, bearers of a particular sensible property have a corresponding sensory appearance in virtue of bearing that property. For example, all red things have a visual appearance—a certain look—in virtue of being red. So, if sounds are audible properties, then their bearers have corresponding auditory appearances. Furthermore, if sounds are audible properties, then they are presumably borne by *non-sounds*—whether objects (Pasnau 1999; Kulvicki 2008; 2014), events (Leddington 2014; 2019), or spatio-temporal regions (as proposed but not defended by Cohen 2010, 306). Thus, if sounds are audible properties, then some non-sounds have auditory appearances. Conversely, if *no* non-sounds have auditory appearances—that is, if *only* sounds have auditory appearances—then sounds cannot be properties; they must be individuals.

In sum, then, we have:

- a. Kivy's Restriction is true if and only if *only* sounds have auditory appearances.
- b. If only sounds have auditory appearances, then sounds are individuals.

---

<sup>4</sup> I'm assuming that these options exhaust the possibilities. Supposing that the world consists of individuals with properties, then, given that sounds are perceivable, they must either be individuals with sensible properties *or* sensible properties themselves.

To resist Kivy's Restriction we must above all resist the idea that only sounds have auditory appearances. There are two ways to do this. First, we might accept that sounds are individuals but insist that some non-sounds nevertheless have auditory appearances. Second, we might reject the view that sounds are individuals in favor of a view of sounds as sensible properties. In Part II, I consider both of these strategies and offer independent reasons in favor of the latter.

### II. Resisting Kivy's Restriction

To resist Kivy's Restriction, an individualist about sounds must allow that some non-sounds have auditory appearances. Here four contemporary views seem promising.

First, according to *Parthood*, sounds are events—disturbings of media—that are proper parts of their medium-involving event sources (see O'Callaghan 2007; 2009b; 2010; 2011). For example, when a baseball collides with a bat, the sound (the disturbing of the air) is a proper part of its source (the ball-bat-collision-in-air). Given that sounds have audible features, Parthood entails that some non-sounds—namely, sound sources—have “parts with audible features,” and so, auditory appearances (O'Callaghan 2011, 396). According to O'Callaghan, this allows us to explain an important phenomenological datum: namely, that we can hear sound sources *along with* their sounds in such a way that “the audible source and the audible sound are not simply phenomenologically unified—they share an audible appearance” (2011, 397–8). Thus, Parthood seems to permit some non-sounds to appear auditorily.

Second, *Identity* holds that sounds are identical with their event sources (see Casati and Dokic (1994) and Casati et al. (2013)). When the ball collides with the bat, the collision *is* the sound, and so, a bearer of audible features such as pitch and loudness. This means that ordinary noisy events such as collisions have auditory appearances. Strictly speaking, this view is compatible with Kivy's Restriction, since the Identity theorist can hold that all and only sounds have auditory appearances. At the same time, by identifying sounds with their ordinary event sources, Identity locates auditory appearances in a manner at odds with the spirit of Kivy's Restriction. Contra Kivy, Identity holds that any noisy event has auditory appearances, and so, is a possible object of musical picturing. In this respect, the Identity theorist can officially embrace Kivy's

## Sonic Pictures

Restriction even while massively extending the scope of what might be auditorily—and so, musically—pictured.

Third, according to *Plurality*, audible qualities such as pitch and loudness can be borne by both noisy everyday events (collisions, etc.) and sounds understood as “pure audibilia”—that is, as metaphysical individuals somehow independent of events and objects that we can see or touch (Soteriou 2018). By denying that sounds alone bear audible qualities—and so, auditory appearances—*Plurality* is straightforwardly incompatible with Kivy’s Restriction.

Finally, according to *Abstracta*, sounds are repeatable, or *abstract*, individuals (Nudds 2001, 221–2; Martin 2012, 345–6). On this view, sounds resemble properties (or universals) in virtue the fact that they can be spatially and temporally “multiply located” (Martin 2012, 345). The idea is that sounds are more like letters or words than ordinary material objects or events: just as the inscription ‘tick tick tick’ repeats the *same linguistic individual*—namely, the word ‘tick’—the ticking clock repeats the *same sonic individual*—namely, the ticking sound—at regular intervals. Proponents of this view argue that, despite being property-like, sounds are *not* properties especially because: (a) we talk about sounds as if they’re individuals; and (b) unlike properties, sounds are in various ways separable from their material sources (Nudds 2001; 2010; Martin 2012). However, as discussed in Part I, the idea that sounds are separable from, or “float off” their material sources readily suggests that auditory appearances belong to sounds alone (Martin 2012, 334). Nevertheless, Matthew Nudds has recently argued at length that, even someone who endorses *Abstracta* can—and *should*—accept that sometimes “material events...themselves are apparent to us in auditory experience” (2014, 482). If his argument is successful, then even *Abstracta* might provide the means with which to resist Kivy’s Restriction.

All four of these views—Parthood, Identity, Plurality, and *Abstracta*—insist that sounds are individuals. To this extent, they remain compatible with Kivy’s Restriction; but inasmuch as they more widely distribute auditory appearances, they provide means to resist it. Even among philosophers who treat sounds as individuals, there is a growing consensus that, in one way or another, ordinary noisy events such as collisions can genuinely appear in auditory consciousness. In this respect, the trend is to say that what we hear, strictly speaking, are not just sounds, but their sources—the ordinary noisy

events that populate our surroundings. This suggests a more direct route to resisting Kivy's Restriction.

*Property* holds that sounds are audible properties of their event sources (Leddington 2014; 2019). Noisy events such as collisions and vibrations can be described as bearing audible qualities such as pitch, timbre, and loudness (Casati, Di Bona, and Dokic 2013; Soteriou 2018). According to *Property*, those audible qualities *constitute* the sound of the event; in other words, *sounds are event-borne audible-quality complexes*. This keeps with philosophical tradition by classing sounds alongside colors among the sensible qualities. However, tradition also treats sounds (like canonical colors) as properties of *objects* (Locke [1690] 1975, II, viii, 14; Pasnau 1999; Kulvicki 2008).<sup>5</sup> So, by taking sounds for properties of *events*, *Property* breaks with tradition.<sup>6</sup> In any case, because an object's perceptible qualities do not mediate our perceptual contact with it, *Property* has the following critical consequence: just as we see objects *in* (but not *by* or *in virtue of*) seeing their colors, so we hear the event sources of sounds *in* (but not *by* or *in virtue of*) hearing their sounds.<sup>7</sup> And just as the primary objects of vision are not colors per se, but color-bearing objects, the primary objects of hearing are not sounds per se, but sound-bearing events. In other words, according to *Property*, we never hear mere noise, only noisy *events*, and the primary bearers of auditory appearances are not sounds, but their event sources.

So far, then, we have five ways to resist Kivy's Restriction (see *Table 1*). These views differ significantly, and we should expect that they will have different consequences for the scope and nature of auditory picturing. But which should we prefer? I have argued elsewhere that considerations of theoretical simplicity—both ontological and syntactic—

---

<sup>5</sup> Why *canonical* colors? Because arguably some events, such as explosions and other light-emitters, have colors, too. Indeed, according to *Property*, if we're going to think of sounds by analogy with colors, we should do them by analogy with the colors of events, so that the sound of an explosion is the auditory analog of its color (Leddington 2019, 625).

<sup>6</sup> And for good reason. The view that sounds are properties of objects is *prima facie* implausible. Most sounds have temporal profiles that differ substantially from those of everyday objects and their properties. The bell that you ring both predates and outlasts the sound of its ringing; the bell does not, however, predate or outlast its color, shape, or size. This is a reflection of the fact that sounds are event-like, not object-like. Arguably, then, if sounds are sensible properties, they are sensible properties of events, not of objects.

<sup>7</sup> On these uses of the phrases 'in', 'by', and 'in virtue of', see Leddington (2014, 323–5).

## Sonic Pictures

give us reason to prefer Property to both Parthood and Identity (2019). I think similar considerations tell against Plurality and Abstracta. Like Property, Plurality recognizes that the primary objects of hearing include ordinary events bearing auditory qualities such as pitch, timbre, and loudness. But while Property stops there, Plurality also insists, along with Abstracta, on the existence of “pure audibilia”—that is, objects of audition that float free of material sources. These, they say, are “sounds.” But I see no reason to recognize the existence of such things, and, other things being equal, we should prefer a view that dispenses with them. Of course, proponents of pure audibilia argue that we nevertheless have reason to recognize the existence of such objects. While fully adjudicating this issue is beyond the scope of this paper, I turn in Part III to one of the more important arguments that have led philosophers to this separability thesis: Roger Scruton’s argument from acousmatic experience.

Table 1.

View?	Parthood	Identity	Plurality	Abstracta	Property
Primary objects of hearing are...	Sounds	Sounds	Sounds <i>or</i> Events bearing audible qualities (pitch, etc.)	Sounds	Events bearing audible qualities (pitch, etc.)
Sounds are...	Proper parts of their event sources	Identical with their event sources	Pure audibilia (abstract individuals)	Pure audibilia (abstract individuals)	Audible-quality complexes borne by events

### III. Scruton’s Argument from Acousmatic Experience

When discussing whether music can represent everyday objects and events such as a “quarrel” or a “forest fire,” Scruton writes that

a sound is [an individual] detachable from the object that emits it, and...is capable of independent existence. However, it is the sound, not the object, that ‘appears’ in the music. How, then, can the music represent the subjects [such as quarrels] that are normally ascribed to it, when those subjects *share no appearance with the music?* (1976, 277–8, my emphasis)

## Sonic Pictures

According to Scruton, it cannot. As we've seen, Kivy's answer is more circumspect: the music cannot represent such things *pictorially*, only *structurally*. Yet beneath this disagreement lies a shared view of the nature of sounds as self-standing objects of perception and the sole bearers of auditory appearances.<sup>8</sup> Despite their many differences, this is a significant alignment and the source of Kivy's conviction that sounds alone can be musically pictured. My aim in this section is to undercut the temptation to this view, which, I think, is easily felt when reflecting on the experience of listening to music.

At the heart of Scruton's philosophy of music is the idea that it is possible to perceptually attend to sounds without attending to their sources. Such "abstract listening" yields what he calls "acousmatic" experience (Scruton 1997, chap. 1; 2009; 2010). Along with Casey O'Callaghan (2011, 380) and Michael Martin (2012, 333–4), Scruton thinks the possibility of acousmatic experience suggests an independence of sounds and sources incompatible with treating sounds as properties. He writes:

Sounds can be detached completely from their source, as by radio or gramophone, and listened to in isolation. This experience—the "acousmatic" experience of sound—removes nothing that is essential to the sound as an object of attention. The striking thing is that sounds, thus emancipated from their causes, are experienced as independent but related objects, which form coherent complexes with boundaries and simultaneities, parts and wholes. (2009, 58)

Now, if particular sounds *can* "be detached completely" from their sources, then it is hard to see how they could be sensible properties of their sources, as I suggest. Sensible properties cannot be detached from their bearers.<sup>9</sup> Thus, if particular sounds "can be detached completely" from their sources, then my view is in trouble.

Yet it seems odd to say that listening to a recording of a musical performance involves listening to sounds that have been "detached" or "emancipated" from their sources. The fact that you can capture the sound of a musical performance and play it back no more shows that sounds can be detached from their sources than the fact that

---

<sup>8</sup> It's worth noting that I do *not* mean that Kivy would accept Scruton's hefty metaphysical commitment to sounds as non-physical "pure events" (2009). Kivy seems wary of such extravagant metaphysical claims.

<sup>9</sup> This requires clarification. Sensible property *instances* cannot be detached from their bearers. Of course, *qua universals*, sensible properties are separable from particular property bearers. But this is not the sort of separability Scruton has in mind. He believes that particular sound *instances* are separable from their sources.

## Sonic Pictures

you can capture the color and shape of a flower in a photograph shows that color and shape can be detached from the flower.<sup>10</sup> Still, it may be easier to listen acousmatically when hearing music played over an audio system than when witnessing a live performance. On the other hand, listening to a musical recording hardly guarantees an acousmatic experience. An audiophile or sound engineer might attend to the music and, at the same time, to how the speakers sound. A guitarist might marvel at the warmth of Wes Montgomery's guitar. And if you listen to Glenn Gould play Bach's *Goldberg Variations*, you might find yourself annoyingly distracted by his humming—hardly an acousmatic experience.<sup>11</sup> So, hearing a musical recording played over an audio system does not prevent the listener from attending to the sources of the sounds that she hears. Nevertheless, Scruton's main point seems to be that, when we *do* listen acousmatically, we experience sounds as individuals that "form coherent complexes with boundaries and simultaneities, parts and wholes," and I think that there is something to this; it does not, however, tell against Property.

Here is why: exactly the same sort of experience is possible not only with color, but also with shape. Take color. Certain forms of abstract visual art, such as color field painting and abstract film, encourage the viewer to visually attend to expanses of color without attending to their bearer(s).<sup>12</sup> In viewing such work, colors may seem to float free of the objects that bear them and appear as individuals that "form coherent complexes with boundaries and simultaneities, parts and wholes." This does not, of course, tell against the view that colors are properties rather than individuals. That something can be an independent object of attention is no reason to think that it is an object. And that

---

<sup>10</sup> The notion of "capture" here should be taken loosely. In particular, it should not be taken to require the sort of "matching" of color and shape criticized by Ernst Gombrich ([1960] 2000). Arguably, the key task for a theory of depiction is to explain how appearances can be captured without being "matched."

<sup>11</sup> This introduces the question whether audio recordings are *transparent*, such that, in listening to a recording of Gould humming, you can be said literally to *hear Gould hum*, or whether, strictly speaking, you hear only the recording. For what it's worth, I do not think that this question can be answered independently of particular pragmatic circumstances. Furthermore, I think that any general argument against the transparency of audio recordings (or photographs and videos) would effectively beg the question by employing an objectionably ad hoc condition on "genuine perception." (Compare Walton (1984 esp. notes 11 and 15) and Martin (2012, pt. 4).)

<sup>12</sup> Consider, for instance, the "visual music" of Oskar Fischinger, some of which can be experienced online thanks to the Center for Visual Music: <http://www.centerforvisualmusic.org/Fischinger/>.

you can attend to A without attending to B does not show that A is independent of or separable from B; it shows only that A and B are nonidentical—and of course a property instance is not identical to its bearer. So, too, consider shape. In a well-known technique for teaching figure drawing, students are asked to attend to the shapes of objects while ignoring their identity, as it is easier to draw the shape of a nose than it is to draw a nose. Students accomplish this by attending to the “negative space” defined by the contours of the nose, and so, attentionally (but only attentionally!) detach the shape of the nose from the nose itself.

In sum, then, the possibility of acousmatic experience does not tell against the property view of sounds. Property holds that sounds are distinct (though not separate or separable) from their sources, and this is sufficient to underwrite the possibility of acousmatic experience.

However, even if the possibility of acousmatic experience does not tell against the Property View, there appears to be an important difference between abstract listening and abstract visual experience of color or shape. As Kivy notes in *Music Alone*, it seems significantly easier to attentionally abstract a sound from its source than a color or shape from its bearer (1990, chap. 1; cf. O’Callaghan 2020, sec. 4.1). This demands explanation. Possibly it indicates that sounds are individuals rather than properties, as attentionally isolating a particular is plausibly easier than attentionally isolating a property. But there is an alternative. We can explain the relative ease of abstract listening without appeal to a particular metaphysics of sounds—namely, as a consequence of *the relative epistemic poverty of audition*. The relative epistemic poverty of audition consists in the fact that, typically, what we know on a purely auditory basis about what we hear is substantially less than what we know on a purely visual basis about what we see.<sup>13</sup> In particular, hearing alone often leaves us unsure exactly what we have heard, which would seem to make it easier to—in Martin Heidegger’s words—“listen away from things, divert our

---

<sup>13</sup> The idea is that hearing is *generally* epistemically impoverished relative to vision. Of course, there are cases where hearing can tell us things that sight cannot. This is trivially true in the dark, but there are plenty of cases where hearing allows us to access things we cannot see. Looking at the wall will generally not tell you where the studs lie; better to knock and listen for the hollow between them. Car horns are useful in part because drivers have blind spots. Typically, however, sight provides much richer information about our surroundings than hearing does. Thus, you will be less likely to miss when driving the nail if you can see the stud; and you may need to turn your head to see exactly where the honking car is located.

ear from them, i.e., listen abstractly” (1977, 52). On this view, the epistemology of hearing, not the metaphysics of sounds, explains why “the ear, far more than the eye, is capable of sustained perceiving in an abstract, non-interpretive mode” (Kivy 1990, 5).

But how *exactly* does relative epistemic poverty entail relative ease of abstract perceiving? My hypothesis is that vision’s epistemic richness imposes a cognitive—and attentional—burden that interferes with abstract perception of properties such as color and shape. In ordinary circumstances, when you open your eyes, not only is there is a great deal *to* notice, there is a great deal that you involuntarily *do* notice. For example, in looking at the tomato before you, you visually notice not only that it is a tomato with a certain color, but also that it has a particular shape and size, that it is lying on a particular surface, casting a certain shadow, and so on. You also visually notice a great deal about the surrounding space. You notice all of this as a matter of course, without any particular effort, and you cannot help but notice most of it. The fundamental *passivity* of most perceptual recognition is frequently overlooked, but helps to explain the relative difficulty of abstract seeing. The bare visual presence of the tomato passively activates a host of visual recognitive capacities; the result is that you automatically see the redness of the tomato as the color of a particular item in a complex scene. This way of experiencing the color of the tomato is foisted on you, and it will *compete* with any attempt to see that color instance abstractly. The result is a kind of cognitive tug-of-war between involuntary visual recognition and attentional abstraction.

If this hypothesis is correct, then it should be possible to facilitate attentional abstraction by reducing the number of perceptual recognitive capacities passively activated by an experience. Indeed, as it turns out, attentionally abstracting the color of the tomato seems easier if it takes up most of your field of view, or if you visually simplify the rest of the scene. For example, you might place the tomato alone on a white tabletop, or (better) deploy a visually neutral screen to obscure everything but a blemish-free patch of tomato-surface. Apparently, such strategies work because they reduce cognitive load (abstract perception is cognitively demanding) *and* the degree to which we cannot help but see the redness *as the color of one particular object* among others.<sup>14</sup>

---

<sup>14</sup> This explains why it seems easier to attentionally abstract the color of a blank wall than that of an object in a complex scene. Thus, too, the exercise from the famous instructional manual, *Drawing on the Right Side*

## Sonic Pictures

The phenomenon cuts across sense-modalities. Some concert-goers shut their eyes to better attend to the “music alone.” Reducing extraneous perceptual recognition facilitates the intended attentional abstraction. Here, the key point is that, in normal cases of pure auditory experience, there is relatively little to notice. Arguably, on shutting your eyes in the concert hall, you cannot help but auditorily notice that a piano is being played somewhere in front of you; but, provided that the hall is quiet, this hardly compares to the wealth of perceptual noticing that automatically accompanies an ordinary visual experience. The vast majority of what is perceptually noticed in a purely auditory experience of a musical performance is *about the sound itself*, and this is all we need to explain the relative ease of abstract listening.

While the details of this account require further development, this is just *one* approach to explaining the relative ease of abstract listening by appealing to epistemic features of audition rather than to the metaphysics of sounds. In fact, Kivy himself offers another such explanation, based on the idea that, thanks to its epistemic richness, vision has greater survival value than hearing (1990: Ch. 1).<sup>15</sup> Whether you accept my argument or prefer Kivy’s explanation is irrelevant. The key is that there is no easy inference from the relative ease of abstract listening to the claim that sounds are individuals rather than properties.

In sum, then, the phenomenon of acousmatic experience gives us no reason to accept that objects of hearing can float free of material processes as both Plurality and Abstracta require. And while I cannot fully settle this issue here—there are other

---

*of the Brain* (Edwards 2012). The beginner is asked to copy Picasso’s drawing of Stravinsky both upside-down and right-side-up. Copying it upside-down yields a much better result, apparently because looking at the Picasso upside-down interferes with perceptual recognition, and so, facilitates attentional abstraction of shape. The next step is learning the technique described above: focusing attention on negative space rather than on the objects themselves. This is a way of *deploying attention* that interferes with involuntary perceptual recognition of the object.

<sup>15</sup> Kivy’s primary aim is to explain why “visual music” of the sort mentioned above (note 15) is less successful than auditory music. He thinks this is explained by the relative ease with which we listen abstractly (itself explained in evolutionary terms). I am not convinced by this. I think the relative power of music has more to do with how we subjectively experience even *fully abstracted* color vs. sound. Consider, for example, that musical dissonance seems to have a potency—and so, to call for resolution—in way altogether unmatched by even the most intense color conflict. There is no obvious relationship between this fact and the relative ease of abstract listening.

arguments to consider—Scruton’s argument is one of a class of arguments that seeks to draw radical metaphysical conclusions from features of auditory experience that can be neatly explained in other ways.<sup>16</sup> As such, these arguments are all quite weak. They certainly should not embarrass anyone who aspires to give a full account of auditory phenomena by appeal to no more than ordinary noisy events and their audible qualities, just as Property insists.<sup>17</sup>

### IV. From Sound in Sound to Event in Event

In Part II, I discussed five main contemporary theories of hearing and sound: Parthood, Identity, Plurality, Abstracta, and Property. All of these views allow that some non-sounds appear auditorily. Thus, it is fair to say that, since whatever appears auditorily can in principle be sonically pictured, there is a growing consensus that, contra Kivy, some non-sounds can be objects of sonic picturing. Next, I argued that, on the basis of theoretical parsimony, we should prefer Property to the other candidate views, and I defended Property against a well-worn objection from musical listening. In the remainder of this article, I want to explore Property’s consequences for sonic picturing.

Property doesn’t just resist Kivy’s Restriction; it turns it on its head. According to Property, just as colors are properties borne by objects (and sometimes events), sounds are audible-quality complexes borne by ordinary noisy events such as collisions. Moreover, just as color-bearers, but not colors, are the primary objects of vision, so sound-bearing events, but not sounds, are the primary objects of hearing. In this case, then, it will typically be just as inappropriate to say that sonic pictures are pictures of sounds as it is to say that visual pictures are pictures of colors. Instead, just as visual pictures are canonically pictures of objects, so sonic pictures will be pictures of events. Of course, visual pictures of objects also depict visible properties—most obviously, shape and color—but those properties are represented *as properties of the represented objects*. In this respect, the representation of the visible properties is secondary or derivative. For

---

<sup>16</sup> See, for instance, the arguments involving loudspeakers and other supposedly abnormal sound sources in Nudds (2010).

<sup>17</sup> Notably, Identity also insists on this. However, it makes the mistake of identifying sounds with their event sources, rather than with the audible qualities of their event sources. Again, for discussion, see Leddington (2019).

## Sonic Pictures

example, Jan Van Eyck's *Arnolfini Portrait* depicts a woman in a green dress. Does it therefore depict a particular shade of green? Yes, but only secondarily: as the color of a particular dress. Similarly, if Property is correct, then sonic pictures will necessarily depict the sounds of the events that they represent, but those sounds will be represented only secondarily: *as properties of their source events*. So, consider the birdsong cadenza in the second movement of Beethoven's Sixth Symphony. What does it depict? Property suggests that the primary object of depiction—what a properly informed listener *hears in* the music—is not mere bird sounds, but a noisy everyday event: *birds singing*. Does the music *also* depict bird sounds? Sure, but only secondarily: as the sound of birds singing. And note that this account matches Beethoven's intention in giving the movement the programmatic title, "*Szene am Bach*," or "Scene by the Brook," rather than "*Geräusche einer Szene am Bach*," or "Sounds of a Scene by the Brook."

Kivy's late paper on musical picturing is entitled "Sound in Sound." This is how he thinks sonic pictures work: we hear sounds in sounds. The view I recommend differs doubly. When we encounter a sonic picture, what we hear, first and foremost, is an event—for example, the playing of a flute. That we experience this event pictorially means that we *hear in it* another event—for example, the singing of a nightingale. So, it's not "sound in sound," but "event in event."

Here's a nice consequence of this shift. When we watch representational films, we also perceive events in events. More precisely, we see worldly events—often involving people—in projection events on a screen. I propose that, just as we experience the projection event on the screen as a moving visual picture of worldly events, so we experience the amplified playback of the film's audio track as a sonic picture of some of those same events. Thus, we can *see in* the on-screen projection event precisely the same event—a conversation, say—that we *hear in* the playback of the audio track. Against this, you might insist that our auditory experience in this case is not pictorial: we do not *hear in* the playback of the audio track the characters speaking; instead, we just hear voices, the sounds of which we involuntarily associate with the visually depicted goings-on in the film. But I find this implausible and untrue to the phenomenology. Just as the on-screen visual picture is *visibly* not a live scene, the playback of the audio track is *audibly* not a live conversation. Indeed, what Robert Hopkins writes of the visual case is equally true of the auditory case: "We seem to see [and hear!] directly neither the events filmed

nor the events of the story told. We are always, plainly, looking at [and listening to!] pictures, and so our experience of those events is only ever seeing [and hearing!] them in the [visuo-auditory!] image before us" (2008, 153; see also 2010). Naturally, this view needs further development and defense, but note that it provides a very tidy explanation of why we experience the audio and visual components of film presentations to be thoroughly fused: they depict the very same events. Our experience of film is a case of unified, simultaneous seeing- and hearing-in.

Returning to the case of musical picturing, consider one of Kivy's main examples:

The twentieth-century French composer Arthur Honegger (1892–1955) wrote a famous piece for orchestra called *Pacific 231*. A 'Pacific' is a kind of steam railway locomotive, and Honegger's composition represents the sound of the engine starting up, barreling along at top speed, slowing down, and finally coming to rest. (2002: 185)

According to Kivy, we hear in the sound of the music the sound of a locomotive. By contrast, according to Property, a performance of Honegger's composition pictorially represents not just the *sounds* of the locomotive's activity, but also *the activity itself*: "the engine starting up, barreling along at top speed, slowing down, and finally coming to rest." In other words, we hear in one event (the performance) another event (the activity of the locomotive). (We also, of course, hear in it the sounds of the locomotive, but again: only secondarily.) Similar emendations apply to Kivy's other cases. For instance, in "Sound in Sound," he argues that "the Dead March [in Handel's *Saul* is]...a musical picture of a dead march or a dead-march picture," and is in that sense "a musical picture of musical sound" (2012, 157). By contrast, I would describe this as a case where one event—the musical performance of the Dead March in Handel's *Saul*—pictorially represents another event: the musical performance of an Old Testament funeral march.

Now, the Dead March in *Saul* is a clear case of what Kivy calls "aided" musical picturing. Without the accompanying text, we'd never even know that the performance was meant as a representation. Though it's less clear, Kivy thinks the same is probably true of Honegger's *Pacific 231* (2002, 185). But if so, is he right "that unaided pictorial representation in music is, if possible at all, too rare a phenomenon to be counted as belonging to music's repertoire of aesthetic possibilities" (2002, 185)? No—I think this is a mistake, though it's understandable given Kivy's focus on Western classical music.

Thus, I'll wrap up this section by describing what I take to be two relatively unproblematic cases of unaided musical pictures.

*Case one: beatboxing.* Beatboxing is “a vocal percussion practice in hip-hop music,” so-called because it developed “from the desire to imitate beatboxes, the first generation of the drum machine” (Bell 2019, 1967). In one of the earliest recorded examples, beatbox pioneer Doug E. Fresh (“The Human Beatbox”) provides a full beatbox accompaniment to MC Ricky D’s rapped vocals in the song, “La Di Da Di” (1985). Since then, beatboxing has developed into a internationally popular, self-standing musical genre in which artists imitate a wide variety of sounds—musical and non-musical—in improvised musical performances. The key point for present purposes is that, without any accompanying text, beatbox performances readily provoke hearing-in. For example, at the beginning of “La Di Da Di,” it’s nigh impossible not to hear in Doug E. Fresh’s beatboxing the playing of an early electronic drum machine. The artistry of the beatboxer is precisely to induce vivid and surprising experiences of hearing in without having to tell the audience what to expect.

*Case two: cover songs.* Magnus et al. offer the following sufficient condition for being a cover: “A version of a song is a cover when it is recorded or performed by an artist or a group who did not write and compose the song themselves and where there is a prior recording which is accepted as canonical or paradigmatic” (2013, 362). I propose that, in many—and perhaps most—cases, covers are unaided musical pictures. Oftentimes, we can hear the canonical track in the cover, and this constitutes a good bit of our aesthetic interest in it, *even if* the cover is also independently musically interesting. For example, in listening to Stevie Ray Vaughan’s masterful instrumental cover of Jimi Hendrix’s “Little Wing,” we can hear in it both the guitar and vocals of the canonical Hendrix track, and much of the pleasure we take in listening lies in appreciating Vaughan’s ability to allow us to do so even while producing a very different piece of music. Moreover, in most cases, properly appreciating a cover aesthetically requires the sort of familiarity with the canonical track that allows you to recognize the cover *as* a cover without aid of words or titles. Just *knowing* that it’s a cover is not enough. If you can’t hear the canonical version in the cover—that is, if all you hear are the cover’s surface features—then you are missing out on something aesthetically essential about the work.

---

## Sonic Pictures

Despite Kivy's attention to it, musical—and, more generally, sonic—picturing has been neglected by philosophers.<sup>18</sup> This is unfortunate, and not merely because depiction has become central to the field of philosophical aesthetics. If I'm right, then sonic picturing is both more common and more aesthetically important than Kivy ever allowed. Indeed, appreciating and understanding the possibility of sonic picturing seems critical to understanding a variety of important auditory phenomena—musical and otherwise.<sup>19</sup>

### Works Cited

- Bell, Emily A. 2019. "Singing and Vocal Practices." In *The SAGE International Encyclopedia of Music and Culture*, edited by Janet Sturman, 1961–1968. Thousand Oaks, Calif: SAGE Publications.
- Berkeley, George. (1713) 1992. "Three Dialogues between Hylas and Philonous." In *Philosophical Works, Including the Works on Vision*. London: J. M. Dent & Sons Ltd.
- Cohen, Jonathan. 2010. "Sounds and Temporality." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman, 5:303–20. Oxford: Oxford University Press.
- Davies, Stephen. 1993. "Representation in Music." *Journal of Aesthetic Education* 27 (1).
- Edwards, Betty. 2012. *Drawing on the Right Side of the Brain*. 4th ed. New York: Tarcher/Penguin.
- Fresh, Doug E., and MC Ricky D. 1985. *The Show & La Di Da Di*. Reality/Fantasy Records.
- Gombrich, Ernst Hans. (1960) 2000. *Art and Illusion: A Study in the Psychology of Pictorial Representation*. Millennium ed., with A new preface by the author. Bollingen Series, 35. The A. W. Mellon Lectures in the Fine Arts 5. Princeton: Princeton University Press.
- Hanslick, Eduard. 1986. *On the Musically Beautiful: A Contribution toward the Revision of Of the Aesthetics of Music*. Translated by Payzant Geoffrey. Indianapolis: Hackett.
- Heidegger, Martin. 1977. "The Origin of the Work of Art." In *Basic Writings*, edited by D F Krell, 143–212. Harper Collins.
- Hopkins, Robert. 1998. *Picture, Image, and Experience: A Philosophical Inquiry*. Cambridge: Cambridge University Press.
- . 2008. "What Do We See in Film?" *The Journal of Aesthetics and Art Criticism* 66 (2): 149–159.
- . 2010. "Moving Because Pictures? Illusion and the Emotional Power of Film." *Midwest Studies In Philosophy* 34 (1): 200–218. <https://doi.org/10.1111/j.1475-4975.2010.00208.x>.

---

<sup>18</sup> The exception is some recent work on whether audio recordings are sonic images (Kulvicki 2006; Martin 2012). I've ignored this debate in this article, choosing instead to focus on broader questions. That said, my position on the issue is implicit in my discussion of film audio in Part IV.

<sup>19</sup> This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 795393.

## Sonic Pictures

- Kivy, Peter. 1984. *Sound and Semblance*. Reflections on Musical Representation. Princeton, NJ: Princeton University Press.
- . 1988. *Ossin's Rage: Philosophical Reflections on Opera*. Princeton, NJ: Princeton University Press.
- . 1990. *Music Alone*. Ithaca, NY: Cornell University Press.
- . 2002. *Introduction to a Philosophy of Music*. Oxford: Oxford University Press.
- . 2011. *Antithetical Arts: On the Ancient Quarrel Between Literature and Music*. Oxford: Oxford University Press.
- . 2012. "Sound in Sound: On Sonic Representation in Eighteenth-Century Music." In *Sounding Off: Eleven Essays in the Philosophy of Music*, 147–163. Oxford: Oxford University Press.
- Kulvicki, John. 2006. *On Images: Their Structure and Content*. Oxford ; New York: Clarendon.
- . 2008. "The Nature of Noise." *Philosophers' Imprint* 8 (11): 1–16.
- . 2014. "Sound Stimulants: Defending the Stable Disposition View." In *Perception and Its Modalities*, edited by D Stokes, S Biggs, and M Matthen, 205–21. Oxford: Oxford University Press.
- Leddington, Jason. 2014. "What We Hear." In *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*, edited by Richard Brown, 321–34. Dordrecht: Springer.
- . 2019. "Sounds Fully Simplified." *Analysis*. <https://doi.org/10.1093/analys/any075>.
- Magnus, Cristyn, P.D. Magnus, and Christy Mag Uidhir. 2013. "Judging Covers: Judging Covers." *The Journal of Aesthetics and Art Criticism* 71 (4): 361–70. <https://doi.org/10.1111/jaac.12034>.
- Martin, M. G. F. 2012. "Sounds and Images." *The British Journal of Aesthetics* 52 (4): 331–51. <https://doi.org/10.1093/aesthj/ays036>.
- Nudds, Matthew. 2001. "Experiencing the Production of Sounds." *European Journal of Philosophy* 9 (2): 210–29.
- . 2010. "What Sounds Are." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman, 5:279–302. Oxford: Oxford University Press.
- . 2014. "Auditory Appearances." *Ratio* 27 (4): 462–82. <https://doi.org/10.1111/rati.12083>.
- O'Callaghan, Casey. 2007. *Sounds: A Philosophical Theory*. Oxford: Oxford University Press.
- . 2008. "Seeing What You Hear: Cross-Modal Illusions and Perception." *Philosophical Issues* 18 (1): 316–38.
- . 2009a. "Sounds." In *Oxford Companion to Consciousness*, edited by Timothy Bayne, Axel Cleeremans, and Patrick Wilken. Oxford: Oxford University Press.
- . 2009b. "Sounds and Events." In *Sounds and Perception: New Philosophical Essays*, edited by Matthew Nudds and Casey O'Callaghan, 26–49. Oxford: Oxford University Press.
- . 2010. "Constructing a Theory of Sounds." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman, 5:247–70. Oxford: Oxford University Press. <http://ocallaghan.rice.edu/research/papers/Constructing.pdf>.
- . 2011. "Hearing Properties, Effects or Parts?" *Proceedings of the Aristotelian Society* 111 (3): 375–405. <https://doi.org/10.1111/j.1467-9264.2011.00315.x>.
- . 2020. "Auditory Perception." In *Stanford Encyclopedia of Philosophy*, Summer 2020

## Sonic Pictures

- Edition. <https://plato.stanford.edu/archives/sum2020/entries/perception-auditory/>.
- Pasnau, Robert. 1999. "What Is Sound?" *The Philosophical Quarterly* 49 (196): 309–24.
- Robinson, Jenefer. 1987. "Music as a Representational Art." In *What Is Music?: An Introduction to the Philosophy of Music*, edited by Philip Alperson. New York: Haven Press.
- Scruton, Roger. 1976. "Representation in Music." *Philosophy* 51 (197): 273–87.
- . 1997. *The Aesthetics of Music*. Oxford: Oxford University Press.
- . 2009. "Sounds as Secondary Objects and Pure Events." In *Sounds and Perception: New Philosophical Essays*, edited by Matthew Nudds and Casey O'Callaghan, 50–68. Oxford: Oxford University Press.
- . 2010. "Hearing Sounds." In *Oxford Studies in Metaphysics*, edited by Dean Zimmerman, 271–78. Oxford University Press.
- Soteriou, Matthew. 2018. "Sound and Illusion." In *Perceptual Ephemera*, edited by Thomas Crowther and Clare Mac Cumhaill, 31–49. Oxford: Oxford University Press.
- Strawson, P. F. 1959. *Individuals: An Essay in Descriptive Metaphysics*. London: Routledge.
- Walton, Kendall L. 1984. "Transparent Pictures: On the Nature of Photographic Realism." *Critical Inquiry* 11 (2): 246–77.
- Wollheim, Richard. 1980. "Seeing-as, Seeing-in, and Pictorial Representation." In *Art and Its Objects*, 2nd ed., 205–226. Cambridge: Cambridge University Press.
- . 2003. "In Defense of Seeing-In." In *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*, edited by Heiko Hecht, Robert Schwartz, and Margaret Atherton, 3–15. Cambridge, Mass: MIT Press.