Sounds and Temporality

Jonathan Cohen*

Space and time are the forms of combination in intuition and serve for applying the categories in concreto (Kant, *Notes on Metaphysics*, AA 18 §5934, 1783–1784).

What is the relationship between sounds and time? More specifically, is there something essentially or distinctively temporal about sounds that distinguishes them from, say, colors, shapes, odors, tastes, or other sensible qualities? And just what might this distinctive relation to time consist in?

Apart from their independent interest, these issues have a number of important philosophical repercussions. First, if sounds are temporal in a way that other sensible qualities are not, then this would mean that standard lists of paradigm secondary qualities offered by Locke, Galileo, and other modern philosophers — lists which include colors, odors and sounds without any significant distinctions — overlook significant metaphysical differences. This, in turn, would threaten to undermine the coherence of the modern understanding of secondary qualities itself. Moreover, a number of authors have recently urged that the essential temporality of sounds makes it impossible to understand sounds as properties (except on a trope theory of properties; see note 3). If true, and given the more or less universal view that colors are properties, this last conclusion would make potentially inapplicable to sounds much of the comparatively well-developed philosophical taxonomy and apparatus that has arisen in philosophical disputes over the status of colors (for presentations of this taxonomy and apparatus see, for example, Byrne and Hilbert (2003); Cohen (2008b)).[1] Therefore, the conclusion that sounds are distinctively temporal would be a serious blow to hopes for a theoretically unified treatment of the sensory qualities.[2] For all these reasons, quite a lot seems to hang on the question of the temporality of sounds.

*Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

[1]Nothing I say here will takes sides in those disputes about color.

[2]Defenders of the essential temporality of sounds may welcome these consequences as overdue correctives to a history of unreflective, "visuocentric" attempts to force sounds into terms taken from debates about color (which has, of course, dominated discussion in philosophy of perception for centuries). Whether the charge of visuocentrism is justified (and whether visuocentrism is a bad thing) depends, of course, on whether sounds really are distinct in the way proposed. But it strikes me as good methodology to start by attempting to leverage tools that have worked in the past; if so, then visuocentrism might make for a useful working methodology even if it turns out to be untenable at the end of the day.

The thesis of this paper is that the contemplated form of sound exceptionalism based on the allegedly distinctive temporal features of sound is unjustified. I allow, of course, that sounds occur in time, that they have temporal durations, and perhaps bear many other interesting relations to time; but I'll suggest that, claims to the contrary in the literature notwithstanding, there is no metaphysically significant disanalogy between sounds and colors on this score. I'll begin by reviewing critically a series of arguments intended to show that sounds are unlike other sensible qualities in that the former must be understood as concrete individuals (§1). Next I'll consider whether the temporal features of sounds — in particular, their having temporal locations and durations — distinguishes them from other sensible qualities (§2). Finally, I'll assess the claim that sounds can be distinguished from other sensible qualities by their capacity to survive qualitative change (§3). I'll conclude that none of these considerations reveal a distinctively temporal aspect to sounds.

1 Sounds as Individuals

The idea that sounds are distinctively and essentially temporal is largely tied up with the contention — pressed especially by recent writers who argue that sounds should be understood as events — that sounds are (unlike colors) concrete particular individuals rather than abstracta such as properties (O'Callaghan, 2007, 17–19).[3] And this has been thought to lead to the conclusion that sounds are temporal in a way that distinguishes them from non-sound sensible qualities such as color because it is generally thought that concrete individuals can, while abstracta (such as properties, and, in particular, color properties) cannot, have temporal locations and extents. I'll return to the question of the temporal features of properties in §2. For now, however, I want to ask what reasons there are for supposing that sounds are concrete individuals rather than abstract properties.

A first reason for taking sounds to be concrete individuals involves noting the ways in which we count and quantify over them. As the church bell chimes noon, we hear one sound (the first chime), then another (the second chime), and so on. After the chimes end, we say that we heard every one of them, or that none of them occurred before noontime, and so on. In these respects, then, sounds are like individuals: we count one chair, then a second chair, and so on, and quantify over all the chairs in the room.

Unfortunately, however, this consideration does not appear to discriminate between individuals and properties. For while we can indeed count and quantify over individuals in these ways, it appears that the same is true of proper-

[3] Of course, if properties are understood as tropes, as suggested by Williams (1953); Mertz (1996), then the worries discussed below vanish; for properties would then be nothing other than (certain kinds of) concrete individual particulars. In what follows, however, I'll be assuming a non-trope understanding of properties both for the sake of sustaining interest in this class of arguments and so as to avoid making property views of sound hostage to such a tendentious theory of properties.

ties. Of course, this follows immediately if we understand properties as tropes (see note 3). But it seems true of properties construed as universals as well. To see this, suppose that the church bell has a size, shape, and color, that these are among its properties (nothing hangs on this; the reader who disagrees is invited to substitute alternative properties of the bell she recognizes), and that these properties are universals rather than tropes. Then the size, shape, and color of the bell are properties, and they are different properties (for they come apart in extension). Moreover, they can be counted (I have mentioned exactly three) and quantified over (the properties I have mentioned). I conclude that the present argument does not rule out a property construal of sounds.

A second piece of support for taking sounds to be individuals rather than properties is that we don't ordinarily think of sounds as being exemplified by individuals. If the sound of the bell's chiming were a property, and assuming we have perceptual access to this sound, it must be exemplified by something. The best candidates for the role of property bearer are the source of the sound (the bell) or the medium of transmission between source and perceiver (in this case, air). Yet it sounds odd to say that the church bell *has* the sound; instead we say that the bell *makes* or *produces* the sound. And if anything it sounds worse to say that the medium *has* the sound; instead we say that the medium *transmits* or *carries* the sound. Thus, this argument concludes, since sounds are not exemplified by individuals, they are inapt for being understood as properties.

Now, as Casati and Dokic (2005) point out, this motivation, which turns on the evidence of ordinary English idioms, will be unpersuasive to anyone who thinks that metaphysics can fail to be revealed by ordinary linguistic usage. Even putting aside this concern about the argument, there are other candidates for the role of sound bearer that may be available to the proponent of sounds as properties. To my ear, at least, it does not sound so bad to ascribe sounds to spatio-temporal regions that are occupied by (and so not identical to) portions of the sound-transmitting medium. In particular, the property theorist might suggest that the sound is exemplified by the spatio-temporal region occupied by its source.[4] If so, then the property theorist will be able to meet the concern that she find a bearer for her properties.

A third piece of support for conceiving of sounds as individuals is that sounds themselves seem to exemplify qualities. In particular, sounds exemplify the so-called "auditory qualities" of pitch, timbre, and loudness. Indeed, sounds stand in a network of similarity relations to one another on the basis of the auditory qualities they exemplify. But if, as some have thought, the things that bear properties are all individuals, then this shows that sounds must be individuals too.

Once again, this argument fails because the criterion on which it turns is not unique to individuals. This can be seen clearly from consideration of colors (everyone's canonical example of properties).

[4]This proposal is compatible with the possibility that spatio-temporal region might exemplify multiple sounds; consequently, the occurrence of multiple sounds in one region is no obstacle to the view.

Quite apart from the current controversy, it is plausible that there are higher-order properties — viz., qualities or properties of properties. And in particular, it is often said that colors exemplify (e.g.) hue, saturation, and lightness/brightness qualities. Admittedly, this may be a secondary, derivative use: it is plausible that x's color exemplifies a hue quality H (say) only if x exemplifies H. Whether or not they are derivative, however, ascriptions of hue, saturation, and lightness/brightness qualities to colors are commonplace and apparently can be true. Additionally, it is a standard view that colors can be organized in a network of similarity and difference relations on the basis of the hue, saturation, and lightness qualities they bear.[5] But if colors — archetypal properties for all parties to the dispute — can themselves exemplify qualities and be organized into a network of similarity relations on that basis, then it is no argument against the property view of sounds that the same is true of them.

The arguments we have considered, then, fail to preclude a property understanding of sounds, and therefore fail to show a significant disanalogy in the way that sounds and color properties are related to time.

2 Temporal Location and Temporal Extent

We've been reviewing, and finding fault with, arguments for the conclusion that sounds are concrete individuals rather than abstract properties. And we've been interested in the latter conclusion because it is generally held that concrete individuals can but abstract properties cannot occupy temporal locations and extents, so this conclusion would provide one way of securing the needed conclusion that sounds are more closely bound up with time than other sensible qualities.

However, whatever one makes of the arguments so far, I take it to be clear (even without the aid of subtle metaphysical arguments) that sounds do have temporal locations and extensions. We say that one sound occurs before another (as it might be, that the one commences at 2:00 and the other at 3:00), and that a first lasts longer than a second (as it might be, that the first extends for an hour and the second extends for only a minute).[6] Whether these observations show that sounds are in some way distinctively temporal, of course, turns crucially on whether non-sound sensible qualities (e.g., color properties) are without temporal locations and extensions.

Given that non-sound sensible qualities are generally taken to be properties, there is an obvious reason for thinking that they cannot bear temporal locations and extensions. Namely, barring a fairly heterodox tropist understanding of properties (whose adoption would limit the attractions of whatever we can say for non-sound sensible qualities — see note 3), we will presumably

[5] It is also true that colors can be organized into other similarity and difference networks on the basis of other similarity metrics defined over others of the qualities they bear; see Kuehni (2003).

[6] Part of the motivation for insisting that sounds have temporal extent comes from the view that they survive change over time; more on the latter in §3.

want to say that these qualities are universals or some other sort of abstracta, hence outside — viz., not located in — space and time.

But, on reflection, this contemplated reason for denying temporal locations and durations to non-sound sensible qualities seems to show far too much. In particular, it would seem to apply indiscriminately to all properties, and to show that all of them are without both temporal locations and extents. Indeed, it would also seem to show that all properties are without *spatial* locations and extents, in so far as abstracta are outside both time and space. But surely it is common ground in the present dispute that some properties — colors, as it might be — can have spatial locations and extents. Or, more cautiously, if it is not granted that colors and other properties literally have spatial locations and extents, it will be allowed that we ordinarily ascribe spatial locations and extents to such properties. And now an obvious worry arises: viz., that whatever makes it true (or at least permissible) to say that colors and other properties have spatial locations/extents will apply, *mutatis mutandis*, as an explanation of what makes it true (or at least permissible) to say that sounds and other properties have temporal locations/extents.

In particular, and although there are many variations on this theme, the most obvious universalist-friendly account of our ascriptions of spatial locations/extents to properties would involve the idea that those locations/extents apply in the first instance to the particular concrete instances of properties, and only in a derivative (possibly pragmatically governed) sense to the properties themselves.[7] By way of analogy, then, a property universalist might treat sounds as properties and then account for our ascription of temporal locations/extents to sounds as derivative from the temporal locations/extents of the instances of sounds.[8]

Now, one possible objection to the line of thought just sketched is that, if successful, it would not explain why colors (say) *lack* temporal locations/extensions, or why sounds *lack* spatial locations/extensions. For, since the strategy seems equally applicable to any property with concrete instances — and equally applicable to any such property in both spatial and temporal forms, it could not provide for any asymmetry in the spatial and temporal features of the properties to which it is applied.

However, I suggest that the alleged asymmetries in the spatial and temporal features of the properties at issue are exaggerated. Continuing with the paradigmatic properties used up to now, it seems to me that colors *can* permissibly (perhaps even truly) be ascribed temporal locations/durations. For one salient example, we complain about the excessive duration of redness when stopped at a traffic light. (On my proposal, this is probably best understood as a complaint about the duration of a particular instance of redness.) Likewise, I join the theoretical consensus in thinking that we can quite naturally ascribe

[7]This pragmatic strategy for rescuing the ascriptions of temporal locations/extents to sounds mirrors in some ways the pragmatic strategy offered by Sorenson (2008) for rescuing the ascriptions of spatial locations/extents to sounds on behalf of the wave theory.

[8]Of course this leaves unanswered the question of what are the instances of sounds; but presumably this is a matter a property theory would have to address in any case.

spatial locations/durations to sounds (Pasnau, 1999; Casati and Dokic, 2005; O'Callaghan, 2007; Sorenson, 2008).

To be fair, I think the ways we talk about sounds and colors *is* asymmetric: we more typically ascribe temporal locations/extents to sounds than colors, and we more typically ascribe spatial locations/extents to colors than sounds. And this asymmetry in our talk about sensible qualities is *prima facie* in conflict with my proposed explanation that treats the sensible qualities symmetrically with respect to their spatial and temporal features. But I think this conflict is only apparent, and can be explained without taking the asymmetry in our linguistic practices as reflecting a metaphysical difference between the sensory qualities themselves. In particular, I think we can explain the asymmetry in our linguistic as a result of our differential sensitivity to spatial and temporal inhomogeneities in color, on the one hand, and to spatial and temporal inhomogeneities in sound, on the other.

On the visual side, the evidence suggests strongly that our perceptual systems are more sensitive to spatial than to temporal inhomogeneities in color. For example, we seem to be far better at discerning color differences in patches presented simultaneously at different regions of the visual field than color differences in patches presented successively in the same region of the visual field — this is why our performance in simultaneous color matching tasks is superior to our performance in successive color matching tasks (Newhall *et al.*, 1957; Pérez-Carpinell *et al.*, 1998). In contrast, we seem to have the opposite bias with respect to auditory discrimination: we are more sensitive to temporal than to spatial auditory inhomogeneities. One demonstration of this difference comes from the discrepancy between our abilities to discriminate melodies on the one hand, and our abilities to discriminate chords on the other. Even musically untrained subjects are very good at distinguishing one four note melody — viz., temporal discontinuity in sound — from another when these are played by a single trumpet (as it might be) from a single location in the space around the subject. In contrast, distinguishing corresponding spatial discontinuities in sound is much more difficult. If four trumpet players stand arrayed before the subject and each play simultaneously a single note, they will thereby produce a chord that constitutes a sonic spatial discontinuity. As any music student who has suffered through this sort of ear training will attest, discriminating one such spatial discontinuity from another takes considerable effort and practice.

Thus, we have reason for thinking that our perceptual systems are more sensitive to spatial than to temporal inhomogeneities in color, but more sensitive to temporal than to spatial inhomogeneities in sound. Given this (presumably contingent) fact about our perceptual endowment, it is considerably easier, and so more useful, for creatures like us to notice, think about, and talk about, the spatial distribution of colors and the temporal distribution of sounds rather than vice versa. But if this is so, then we should not take the asymmetry in our ascriptions of spatial and temporal features to sounds and colors to reveal an underlying metaphysical difference between these two sorts of perceptual qualities. Rather, this asymmetry is plausibly a result of our contingent perceptual endowment (and what, given that endowment, it is useful

and interesting for us to talk about).

Returning to the main thread, it seems to me that the abstractness of properties in general, and color properties in particular, does not prevent them from sharing the types of temporal features (namely, temporal locations and durations) that sounds have. In short, then, temporal locations and durations do not make for a distinctively temporal aspect of sounds of the sort we were seeking.

3 Survival Through Qualitative Change

A further reason some have given for thinking sounds are distinctively temporal is based on the allegation that sounds survive qualitative changes in their auditory qualities (e.g., their pitch, loudness, or volume). Thus, O'Callaghan writes that,

> . . . sounds survive changes to their properties and qualities. A sound that begins high-pitched and loud may continue to exist though it changes to being low-pitched and soft. An object does not lose its sound and gain a new one when it goes from being high-pitched to low-pitched, as with an emergency siren's wail. The sound of a word begins with certain audible characteristics and ends with others, but a pitch shift is not the end of a sound. Determinate perceptible or sensible qualities, however, do not survive change in this way. The red color of the fence does not survive the whitewashing. The dank smell of the dog does not survive the perfuming. Particulars, such as the fence and the dog, however, survive changes to their qualities (O'Callaghan (2008a, 5ms); cf. O'Callaghan (2007, 22)).

The charge that sounds survive qualitative changes is not only offered as a way in which sounds are distinctively temporal, but also as yet another reason for resisting a property theory of sounds of the sort held by Pasnau (1999, 2000). Indeed, O'Callaghan goes on following the quoted passage to claim not only that sounds survive qualitative change, but that their temporally evolving pattern of qualitative features is just what individuates one sound from another.

I do not believe these considerations about survival are decisive. The first point to make is that, even if it were true that sounds survive qualitative change, this wouldn't distinguish sounds from colors given a standard (though controversial) description of the phenomenon of color constancy.[9] On this standard description, instances of color constancy involve one color that differs in its qualitative presentation — e.g., the one color has different chromaticities (as revealed by color matching tasks) when presented under each of two different

[9] In particular, I myself reject this description of color constancy for reasons alluded to in note 16 and discussed in much greater detail elsewhere (Cohen, 2008a). But the dialectical point I am making stands: *if* one accepts the standard description of color constancy, then the alleged survival of sound across qualitative change fails to distinguish sounds from colors.

illuminants. The presentations can be simultaneous (as in cases of simultaneous color constancy), or not simultaneous (as in cases of successive color constancy). If this is the right description of the phenomenon, then in instances of color constancy there is a single color that survives qualitative change (at different spatial regions and a single time, in simultaneous cases, or at one spatial region and different times, in successive cases). Consequently, the claim that sounds survive qualitative change would fail to mark out a distinctively temporal dimension of sounds among the sensory qualities.

However, and even more importantly, I am not convinced that sounds survive qualitative change. Rather, it seems to me, the evidence is equivocal between survivalism and non-survivalism, so these should be treated as something like alternative and equally acceptable methods of bookkeeping; of course, this would also mean that it is inappropriate to use a criterion of survival to rule out metaphysical theories of sound that are incompatible with it.

To see why, consider one of O'Callaghan's examples, the wail of the siren, that occurs over an extended temporal interval T. Focus on two instants in T — call them t_0 and t_1 — such that what you hear at t_0 differs in respect of pitch from what you hear at t_1. Grant, as seems plausible, that you hear a sound at t_0 and that you hear a sound at t_1. (This leaves open that there may be other things that you hear at these instants, including possibly other things that you hear in virtue of hearing these sounds). Is it true that the sound you hear at t_0 the same as (numerically identical to) the sound you hear at t_1?

Here are two ways of answering this question that strike me as equally acceptable.

- First answer: survivalism. You hear a single, temporally extended sound over the entire interval T — that is, you hear a single extended sound at each moment that comprises T. *A fortiori*, you hear that single sound at both t_0 and t_1. Now, it is a feature of the hypothesized case that the pitch exemplified at t_0 is distinct from the pitch exemplified at t_1, so it follows from the present answer that the qualities of the one sound heard — in particular, its pitch qualities — change over time. And this means that a sound can survive changes to its pitch qualities. Indeed, analogous responses to slight variants of our set-up that involve variations in timbre and loudness suggest that a sound can survive changes to these auditory qualities as well. In addition to the single sound we hear over T, there are, of course, many temporal parts of the sound. While standards for temporal part individuation can vary, one potentially fruitful strategy would be to individuate the temporal parts of the sound heard over T such parts by their auditory qualities. It would then follow that the temporal parts of the sound, unlike the sound itself, don't survive changes to their auditory qualities. And, of course, the duration of these temporal parts is less than T itself. But the one, changing sound we hear extends in duration over the entire interval T.

- Second answer: non-survivalism. You hear a number of distinct sounds over interval T. These sounds can be distinguished by differences in their

> pitch (or other auditory) qualities. It follows from this that a sound does not survive changes in its auditory qualities. Since the pitch quality exemplified by what you hear at t_0 is distinct from the pitch quality exemplified by what you hear at t_1, the sound you hear at t_0 is numerically distinct from the sound you hear at t_1.[10] In addition to the multiple individual sounds we hear over T, there is, of course, the single, temporally extended stream of which these sounds are temporal parts. It is natural to say that the stream can bear auditory qualities at particular moments over which it extends, and that it inherits these qualities from the temporal parts it has during those particular moments (just as it is natural to say that a chair can bear color qualities at the places over which it extends, and that it inherits these qualities from the spatial parts it has located at just those places). If we do say this, then because its temporal parts differ in their auditory qualities, we should say that the single stream exemplifies auditory qualities over T. And this is just to say that the temporally extended stream, unlike the sounds that are its temporal parts, survives changes in its auditory qualities.

Now, it must be admitted that, as O'Callaghan points out in the quoted passage, some aspects of our ordinary usage favor the survivalist option. We talk about *the sound* (singular) of the siren, not *the sounds* (plural) of the siren, despite the fact that the pitch we hear when we hear the siren at t_0 is distinct from the pitch we hear when we hear the siren at t_1. Similarly, if you instruct someone to reproduce the sound of the siren she will ordinarily produce a temporally extended sound stream that varies in its pitch qualities between t_0 and t_1, thereby answering your request for one sound by producing something — presumably a sound, if your interlocutor is cooperative — that varies in (hence, survives changes in) pitch.

But ordinary usage points the other way as well. We might describe an extended stream produced by a siren in terms of its being made up of one sound (with one stable set of auditory qualities) at t_0, then another (with another stable set of auditory qualities) at t_1, and so on. (Breaking down a stream into these isolated and qualitatively unchanging parts is one natural mode of instructing novices — e.g., children — in the oral reproduction of sound streams of interest.) Similarly, we are inclined to describe the auditory effect of footsteps as a sequence of qualitatively uniform sounds (the individual steps) rather than a single repetitive sound that alternates between distinct qualities (a step, then a silence, then a step, etc.). These ordinary ways of talking about sound, it seems to me, fit better with non-survivalism.

It appears, then, that ordinary usage pulls in both directions, and so does not settle the issue.

One might, therefore, look to semantics as a further possible way of choosing between survivalism and non-survivalism. In particular, one might notice

[10] In order to set aside further issues about the metaphysics of time and change, I assume here that the sounds recognized by this answer have non-zero temporal durations less than the duration of T.

that *sound* can be used as a mass noun, and infer from this that what the word picks out is a collection of individuals (e.g., a surviving stream, as per survivalism) rather than an individual. Evidence that *sound* functions as a mass noun consists in the observations that, in environments such as *sound of a siren*, it cannot be directly modified by a number without specifying or at least implying a unit of measurement (1a–1c), and is more acceptable with a definite than an indefinite article (2a–2c). In these respects, *sound* in this environment patterns with mass nouns such as *furniture* rather than count nouns such as *chair*:

1a Mary hears the sound/?two sounds of a siren.

1b Mary sees the chair/two chairs.

1c Mary sees the furniture/*two furnitures.

2a Mary hears the sound/?a sound of a siren.

2b Mary sees the chair/a chair.

2c Mary sees the furniture/*a furniture.

But these reflections seem unconvincing as well. For one thing, the very same criteria suggest that *sound* functions as a count noun in other environments (e.g., *hear a sound*), as shown by (1d) and (2d).

1d Mary hears the sound/two sounds.

2d Mary hears the sound/a sound.

More fundamentally, there is no reason to think that the semanticist's notions of mass and count correspond to the metaphysical categories of qualitatively-stable individual and qualitatively-unstable stream. Consequently, the evidence we've been considering seems, once again, unable to choose between survivalism and non-survivalism.

Finally, one might hope to support survivalism by pointing to empirical research on the way the auditory system segments or parses the occurrent flux of auditory information into temporally extended streams — a process Bregman (1990) calls "auditory scene analysis." It seems clear that human perceptual systems do perform this sort of analysis: we distinguish auditorily the wail of the siren from the temporally overlapping roar of the passing automobile engine. Moreover, it is clear that the streams that are so segregated survive changes to their auditory qualities; indeed, a main focus of this area of empirical research is to discern the exact dimensions of variation and amount of variation along these dimensions that a single stream can have before it is parsed as more than one stream.[11]

There is a further reason for being interested in the streams picked out by auditory scene analysis: they are plausibly the loci for the binding of auditory

[11] For example, it turns out that streams fail to survive sufficiently large variation in pitch (more accurately: frequency), spatial location, timbre, loudness, and temporal location (Bregman, 1990, chapter 2).

qualities.[12] After all, it is not that audition merely represents loudness at one time, high pitch at another, and so on; were this the case, we would be unable to discriminate a scenario with one loud, high tone and a second soft, low tone, on the one hand, from one with a loud, low tone and a soft, high tone. The best explanation of our capacity to make such discriminations is that we represent auditory qualities not by themselves, but *as applying to individuals*. For if so, we could describe the first scenario as one in which loudness and highness are represented as applying to one individual while softness and lowness to a second individual; this would then differ from the second scenario, under which loudness and lowness are represented as applying to a first individual while softness and highness are represented as applying to a second. This idea connects with auditory streams in so far as the latter are plausible fillers of the role of auditory individuals needed for the success of the proposed explanation of the binding of auditory qualities.[13]

Psychological considerations about auditory scene analysis and auditory binding, then, give us further reason (in case more was needed) for thinking that there are sound streams that survive qualitative change, and that any acceptable philosophical account of sound must recognize them. But what is not clear, and what is not settled by the empirical research, is whether sounds should be identified with such streams (as per survivalism) or not (as per non-survivalism). After all, non-survivalists, too, accept the existence and importance of sound streams. And non-survivalists have at hand a gloss on what auditory scene analysis amounts to; they will characterize auditory scene analysis as a process of grouping or integrating distinct sounds that fail to survive qualitative change, rather than as individuating one qualitatively changing sound from another.[14] Yet again, it seems that the phenomena at issue are neutral between a survivalist and non-survivalist metaphysics of sound.

More generally speaking, I doubt that there are perceptual phenomena — as opposed to theoretically laden descriptions of perceptual phenomena — that are better accounted for by one or the other of these views. This is because survivalists and non-survivalists agree about what there is: they both recognize the existence of temporally extended entities that survive change in auditory qualities (the survivalist treats these entities as sounds, the non-survivalist does not). And they both recognize the existence of temporal parts of such streams, which do not survive change in auditory qualities (the non-survivalist treats these as sounds, the survivalist does not). Finally, they both agree that we have perceptual contact with the temporal parts that make up the streams, and, derivatively, with the streams themselves, so both are in a position to say

[12]The problem under discussion here is an auditory version of the many-properties problem discussed by Jackson (1977), and later by Clark (2000). For discussion and comparison between auditory and visual instances of perceptual binding, see O'Callaghan (2008b).

[13]For persuasive arguments against the alternative view that auditory properties might serve as the needed individuals, see O'Callaghan (2007, 19ff).

[14]For what it is worth, Bregman (1990, chapter 1) himself repeatedly characterizes auditory scene analysis as a process of grouping auditory components.

that we succeed in hearing sounds.[15] This suggests that the two views will not differ in the treatments of perceptual phenomena they make available, but only in their descriptions of the perceptual phenomena — in their understandings of how the explanatory burdens of accounting for such phenomena are distributed across their shared apparatus.[16]

[15]That we can make auditory contact with the temporally extended sound stream despite at every instant only being auditorily exposed to one of its temporal parts is a special case of the so-called puzzle of "presence in absence" discussed by O'Regan and Noë (2002); Noë (2004, 2006b, 2007); Noë finds it generally puzzling that we manage to sense at one time and place (and know that we sense at one time and place) whole objects — a whole tomato, as it might be — given that we seem to make direct perceptual contact with only a small subset of spatial/temporal parts of the whole — just the facing surface of the tomato, as it might be. His solution to the puzzle is to hold that our perceptual contact (alternatively, sometimes, just our "sense" of our perceptual contact) at t with the whole is constituted in part by our implicit grasp of "sensorimotor contingencies" — counterfactuals about what we would sense were we to change our spatiotemporal relation to the whole. Thus, for Noë, perceptual contact with wholes requires that the parts be perceptually *accessible*, not that they be perceptually *accessed*.

A *prima facie* objection is that this treatment is inapplicable to sound stream perception: when you hear a siren wail W (a temporally extended sound stream) at t, past temporal parts of W are *not* accessible (indeed, depending on how the counterfactuals are understood, perhaps necessarily not accessible) in any of the ways you, as an ordinary perceiver, have at your disposal (Clark, 2006, cf.). But Noë (2006a) is unbothered by this worry: he holds that there is no "presence in absence" phenomenology in the auditory case of the sort that motivated his account of the visual case.

I find this response unsatisfactory. Perhaps I am just a phenomenological boor, but I confess to not recognizing in either the visual or the auditory case the alleged phenomenology of "presence in absence" — although of course I recognize that we have a "sense" (viz., a belief) that we make perceptual contact with wholes. If the intended scope of the account is limited to phenomenology, then, I don't see that there's anything to be explained.

In contrast, if we take Noë's view as an account of how we make perceptual contact with wholes, then there is something to be explained, but, as have seen, Noë's account is inapplicable to cases of auditory perception of sound streams, and this should make us doubt it in other (e.g., visual) cases as well. A more plausible and more general view on this matter is that, just as S can make non-perceptual contact with x by achieving non-perceptual contact with only a subset of x's proper parts, so too S can achieve perceptual contact with x at t by virtue of achieving perceptual contact with only some proper parts of x at t (or even merely in some neighborhood of t, in order to avoid the problems about bottom-up atomism discussed by Matthen (2008)). This condition on perceptual contact is, of course, compatible with the perceiver's failing to bear counterfactual relations to currently unsensed spatial or temporal parts of x.

Finally, if the question is why we *believe* that we achieve perceptual contact with wholes despite occurrently contacting only proper parts, then surely the best answer is that we believe this because we *do* make perceptual contact with wholes. Once again, there is no need for perceivers to bear counterfactual relations to currently unsensed spatial or temporal parts.

[16] There is an instructive parallel here with certain issues concerning color constancy. Philosophers (following perceptual psychologists) have traditionally described the phenomenon of color constancy as a kind of invariance in apparent color across changes in illumination, and have then appealed to the phenomenon as the most important empirical motivation for views that identify colors with illumination-independent properties of surfaces. However, as I have pointed out elsewhere (Cohen, 2008a), this description of the phenomenon is either empirically inadequate or theoretically biased. A more neutral, and more empirically adequate, description is that apparent colors exhibit variance (in one sense) and invariance (in some other sense) across changes in illumination. Given this more adequate characterization, it is then up for grabs whether what is variant across such changes or what is invariant across such changes should be understood as color; and therefore it is also up for grabs whether or not the phenomenon of color constancy (so-called) lends support to illumination-independent accounts of color.

I take the foregoing to show that we can think about and talk about both streams and temporal parts of streams, that both answer to at least some parts of the best job description for sounds we know how to formulate, and that no total theory of sound and sound phenomena can do without either. In particular, then, it seems to me that what is known about the (neutrally described) facts about sound survival is agnostic between survival and non-survival views. (I am not asserting that there is no fact of the matter about which is right, but only that if there is a fact of the matter, the sources of evidence I have discussed fail to disclose what it is.) Once again, therefore, considerations about survival through qualitative change fail to provide for a temporal aspect that would distinguish sounds from other sensory qualities.

4 Conclusion

Sounds occur at particular times and have temporal durations. In those senses, and perhaps others, they are creatures of time. But I do not see that the case has been made that there are *distinctively* temporal aspects of sounds — aspects that distinguish them in metaphysically significant ways from other sensory qualities. While it may be that metaphysical differences between sounds on the one hand and colors, tastes, and odors, on the other, ultimately undermine the prospects for a uniform treatment of sensory qualities, it seems to me that what we know about the temporal features of sounds does not license this conclusion at this time.[17]

References

Bregman, A. S. (1990). *Auditory Scene Analysis*. MIT Press, Cambridge, Mass.

Byrne, A. and Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences*, **26**(1), 3–64.

Casati, R. and Dokic, J. (2005). Sounds. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*.

Clark, A. (2000). *A Theory of Sentience*. Oxford University Press, New York.

Clark, A. (2006). That lonesome whistle: A puzzle for the sensorimotor model of perceptual experience. *Analysis*, **66**(1), 22–25.

Cohen, J. (2008a). Color constancy as counterfactual. *Australaian Journal of Philosophy*, **86**(1), 61–92.

Analogously, it begs the question in favor of sound survivalism to say that it is the *sound* rather than (as the non-survivalist maintains) the *sound stream* that persists across changes in auditory qualities.

[17] I am grateful to Kalle Aho, Craig Callender, Casey O'Callaghan, and Chris Wüthrich for helpful discussion of these matters.

Cohen, J. (2008b). *The Red and The Real: An Essay on Color Ontology*. Oxford University Press, Oxford. Forthcoming.

Jackson, F. (1977). *Perception: A Representative Theory*. Cambridge University Press, New York.

Kuehni, R. G. (2003). *Color Space and Its Divisions: Color Order from Antiquity to the present*. Wiley, New York.

Matthen, M. (2008). On the diversity of auditory objects. *European Review of Philosophy*, **7**.

Mertz, D. W. (1996). *Moderate Realism and its Logic*. Yale, New Haven.

Newhall, S. M., Burnham, R. W., and Clark, J. R. (1957). Comparison of successive with simultaneous color matching. *Journal of the Optical Society of America*, **47**(1), 43.

Noë, A. (2004). *Action in Perception*. MIT Press, Cambridge, Massachusetts.

Noë, A. (2006a). Experience of the world in time. *Analysis*, **66**(1), 26–32.

Noë, A. (2006b). Experience without the head. In T. S. Gendler and J. Hawthorne, editors, *Perceptual Experience*, pages 411–433. Oxford University Press, New York.

Noë, A. (2007). Real presence. *Philosophical Topics*, **33**(1), 235–264.

O'Callaghan, C. (2007). *Sounds*. Oxford University Press, Oxford.

O'Callaghan, C. (2008a). Constructing a theory of sounds. *Oxford Studies in Metaphysics*.

O'Callaghan, C. (2008b). Object perception: Vision and audition. *Philosophy Compass*.

O'Regan, J. K. and Noë, A. (2002). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, **24**(5), 883—975.

Pasnau, R. (1999). What is sound? *Philosophical Quarterly*, **49**(196), 309–324.

Pasnau, R. (2000). Sensible qualities: The case of sound. *Journal of the History of Philosophy*, **38**, 27–40.

Pérez-Carpinell, J., Baldoví, R., de Fez, M. D., and Castro, J. (1998). Color memory matching: Time effect and other factors. *Color Research & Application*, **23**(4), 234–247.

Sorenson, R. (2008). Hearing silence: the perception and introspection of absences. In M. Nudds and C. O'Callaghan, editors, *Sounds and Perception: New Philosophical Essays*. Oxford University Press, Oxford.

Williams, D. C. (1953). The elements of being. *Review of Metaphysics*, **7**, 3–18 & 171–192.