I think that we should conceive of sound as like colour, rather than as like light. That is to say, we should continue to treat sound as the object of hearing and we should think of sounds as existing within the object that ‘makes’ them.

— Robert Pasnau (1999, 316)

I take it, however, that we want to attribute the tonal qualities of the noises made to the nature of the thing that produces them, and thus ascribe the tonal characteristics to the thing itself as causal properties.

— J. O. Urmson (1969, 123)

Robert Pasnau rejects the “standard” view that sounds are waves in media like air and water in favor of the view that sounds are properties of objects, like bells, that make the media move (Pasnau 1999, 309). Sounds are perceived to have locations, and those locations seem to correspond to the objects that “make” the sounds. Waves produced by an object fill the space around an observer, and thus are not correctly perceived to be at their source. So, Pasnau suggests, absent an error theory, it is better to identify sounds with objects’ vibratory properties than with pressure waves such objects produce. Casey O’Callaghan (2007, Ch. 3) dismisses the wave view for similar reasons but claims that sounds are events in which objects disturb the media around them. O’Callaghan’s approach relates closely to one proposed by Roberto Casati and Jerome Dokic (1994; 2005), according to whom sounds are vibration events rather than disturbings of media. Pasnau (forthcoming) recently expressed his support for Casati and Dokic’s take on things, while Roy Sorensen (2007, 281–285) recently defended the wave view against Pasnau and O’Callaghan.

This recent interest in sounds is welcome, since they have been relatively ignored over the past century. The growing consensus is that sounds differ dramatically from colors. While colors are qualities, sounds are particulars: either waves or vibratory events. In addition, all of the views just sketched insist that sounds are transient in a way
that colors are not. Sounds are more like movements than like colors. Objects move in many ways, but it rarely makes sense to ask what kinds of movements an object has, as opposed to how an object is moving now or then.

The following suggests that philosophers have overlooked an impressively promising candidate for being sounds. Sounds are stable properties of objects that seem to have them. More specifically, sounds are dispositions of objects to vibrate in response to being stimulated. Sounds are perceived transiently, but they are not perceived as being transient and they are not in fact transient. This conception of sounds—the stable property view—casts them in a role more akin to colors than other theories do.

O’Callaghan believes “Visuocentrism’ has shaped our understanding of perception and its role” (2007, 2) and suggests that colors are a poor guide to understanding sounds. I disagree. The following shows that reflectance physicalist accounts of color (Hilbert 1987, Tye 2000, Byrne and Hilbert 2003) are excellent guides to understanding sounds. Phenomenological and empirical considerations that favor reflectance physicalism—especially those concerning color constancy—have strict analogues in the auditory case that suggest the stable property view is correct. Everyone mentioned above conceives of sounds as physical, in the relevant sense of the word, but none of those views can easily explain the facts about auditory constancy. Furthermore, the stable property view ably handles phenomena that the other accounts highlight in support of their own positions.

The stable property view sits uneasily with many ordinary ways of talking about sound. The Urmson epigraph suggests it is appropriate to ascribe “tonal qualities” to objects in a stable way, but this does not obviously apply to sounds more generally. Few philosophers think that ordinary language should dictate the practice of metaphysics or the study of perception, but grand departures from common expressions ought to be justified by theoretical gains. The following suggests that the departures are not as grand as one might think, and that they are indeed justified.

Section 1 describes and motivates Pasnau’s (1999) view of sounds, because his is closest to the present proposal in treating sounds as properties. Section 2 unpacks the stable property view, while Section 3 considers an objection and ultimately amends the core proposal. Section 4 considers some positive reasons in favor of the stable property view derived from empirical work on auditory constancy. Section 5 considers whether the fact that we often take ourselves to hear events suggests that sounds are events, as Casati, Dokic, O’Callaghan, and Pasnau (forthcoming) suggest they are. Section 6 discusses the role loudness plays in the stable property view. And Section 7 situates this proposal with respect to the problem of identifying any so-called secondary qualities with physical properties of the objects that seem to have them.

1. Pasnau on sound

Objects spin, bounce, rock, and roll. In addition, they vibrate. We are often in a position to see, feel, and hear movements. You see the acrobat spin, hear the train approach, and feel the boat bob in the water. Vibrating is a way an object moves. Pasnau writes:

I propose identifying sound with the vibrations of the object that has the sound. More cautiously, I would say that sounds either are the vibrations of such objects, or supervene on those vibrations. (Pasnau 1999, 316)

According to Pasnau, sound perception is just one of the ways in which we perceive objects’ movements, or at least something that supervenes thereupon. Sounds are not perceived as movements, of course. One need never suspect, on the basis of hearing things, that sounds are movements, and sounds are not conceived of as movements. This should not surprise anyone familiar with philosophical work on color. Few theories of colors suggest that their true natures are made fully manifest by experiences of them (Johnston 1992) and it would be unsurprising if this were true of all the secondary qualities.

One thing that makes sounds similar to other perceived movements
for Pasnau is that they are transitory states of the objects that manifest them. The acrobat ceases spinning, the train stops at the station, and the boat’s bobbing becomes a swaying. Similarly, the bell stops ringing and the crowd quiets down. This feature of sounds distinguishes them from colors. “The key difference is that colours are generally stable, lasting properties, whereas sounds typically last but a moment, and depend on how forcefully one object strikes another” (Pasnau 1999, 322). We do not perceive movements as “stable, lasting properties” of objects, and Pasnau thinks sounds fit this mold as well.

This distinction between sounds and colors helps Pasnau meet one of Berkeley’s objections (Dialogue I, 171–172; Pasnau 1999, 321) to the view that sounds are vibratory properties. In particular, Berkeley claimed that objects make no sounds in the vacuum, but they do vibrate, so it is wrong to identify sounds with those vibrations. Pasnau disagrees: “just as we are inclined to say that objects have their colours in the dark, so should we say that objects make a sound in a vacuum” (Pasnau 1999, 322). The inclination to think objects make no sounds in the vacuum can be traced to the fact that “sounds typically last but a moment.” Pasnau suggests that if objects tended to vibrate audibly, more or less constantly and without the assistance of being struck, then it would be as plausible that objects have sounds in the vacuum as it is that they have colors in the dark. This is because such objects would seem to have sounds as a part of their “intrinsic nature” and they would not seem to lose such qualities merely because of changes in their external circumstances (Pasnau 1999, 322). Even in a world like ours, in which sounds are transient properties, Pasnau thinks the foregoing suggests vibrating objects have sounds in the vacuum.

The next section argues that Pasnau has underestimated some interesting points of analogy between sounds and colors. It is easy to respond to vacuum worries because sounds are in fact “stable, lasting properties” of the objects that seem to have them.

2. Sounds as natural modes of vibration

Pasnau is right to suggest that “we should conceive of sound as like colour, rather than as like light” (1999, 316). With that in mind, consider the following admittedly controversial claims:

C1. Colors are properties of the objects that seem to have them.

C2. Light is the means by which we come to know the colors of objects perceptually.

C3. Without light, objects have colors, but the colors cannot be seen.

S1. Sounds are properties of the objects that seem to have them.

S2. Compression waves are the means by which we come to know the sounds of objects.

S3. Without compression waves, objects have sounds, but the sounds cannot be heard.

Pasnau (1999) could agree with all of these claims, though he says little explicitly concerning C1–C3.° If one accepts S1, as Pasnau does, it is rather natural to accept S2 as well, since that seems to be the role compression waves would have to play on such an account of sounds. The epigraph makes it clear that Pasnau thinks of sounds as being more like colors — properties of objects — than like light — the means by which color is perceived. Pasnau’s treatment of the vacuum reflects his endorsement of S3. Objects have sounds in the vacuum when they vibrate, but they make no compression waves and thus cannot be heard.

It’s tempting to liken the vacuum to the dark, because in that situation the object’s properties are imperceptible, but present, like colors in the dark. But there are two ways in which one can fail to receive compression waves from an object: when it is in the vacuum

---

1. Pasnau (forthcoming) rejects C1 and S1 and argues that sounds and colors are events — i.e., particulars — not properties.
and when it is not vibrating. Pasnau ignores the latter case because he thinks sounds are vibratory properties. Sounds are transient, like vibrations, or so most philosophers have thought. There is another way to understand the analogy between color and sounds, however. Perhaps sounds are not vibrations per se but dispositions to vibrate in response to certain kinds of stimulation. Colors are not the reflecting of light by a surface so much as a surface's disposition to reflect light in a certain manner. An object that doesn't vibrate makes no sound, but that is a far cry from saying it has no sounds.

To motivate this view of sounds, consider the following controversial claims:

Objects give off light when stimulated by light.

Objects in a medium give off compression waves when they are thwacked.

Somehow or other, an object must be thwacked in order for it to give off compression waves, just as objects must be illuminated if they are to give off light. The character of stimulation can vary substantially: bats, fingers, stones, other compression waves, and stiff breezes can all make objects vibrate. A thwack, speaking quite generally, imparts energy to an object. If sounds are dispositions to vibrate when thwacked, then an object in the dark, sound-wise, is an object un-thwacked or an object in the vacuum. The controversial analogy between sounds and colors thus includes:

S4: Without vibrating, objects have sounds, but those sounds cannot be heard.

This completes the analogy by recognizing that there is an interesting disanalogy between the waves that carry information about each. Light requires no medium through which to propagate while compression waves do. The audible dark is slightly more complicated than its visual cousin.

Consider two more uncontroversial points:

An object's color depends on the proportion of light it reflects across the visible spectrum.

An object's sound depends on the way in which it vibrates in response to being thwacked.

The claim about light and color helps motivate the view that colors are dispositions of objects to reflect light in the visible spectrum: reflection physicalism (Hilbert 1987, Byrne and Hilbert 2003). These considerations are inconclusive, but they can play a similarly limited role in motivating the view that sounds are stable dispositions to vibrate at different frequencies in response to stimulation.

Objects are disposed to vibrate in particular ways when thwacked. Interestingly enough, these vibratory dispositions remain stable as modes of stimulation and contexts change. Any elastic object tends to damp out certain frequencies that stimulate it and resonate with others to differing degrees. A good thwack with a hard object, like a bell's clapper, is tantamount to giving an object a dose of white noise — vibrations across the range of audible frequencies in roughly equal proportion — akin to illuminating an object with white light. Some of those frequencies have little if any effect on how the object vibrates

2. An elastic object is one that, if deformed, returns to the state it had before being deformed. Most objects we would call "rigid" — chairs, coffee mugs, windshields, e.g. — are elastic in the sense relevant here. If they were truly rigid, they would not vibrate, since their parts would not move with respect to one another.

3. Technically, a thwack is a short-lived, intense stimulus, which would be mathematically represented by a delta function — a function that has some positive value at some point in time and is zero everywhere else. (This way of putting things harmlessly papers over the distinction between the Dirac delta function and the Kronecker delta function.) The Fourier transform of a delta function — the decomposition of a delta function into its frequency components — is white noise: i.e., a function that has the same intensity across the range of frequencies (Pain 2005, 290). In this way, the un-technical talk of thwacks has decent mathematical credentials.
while others have considerably greater effect. Objects have resonant modes that lead them to vibrate at certain natural frequencies. These resonant modes depend on the structure and material of the object, not the way it is stimulated. Striking an object in different ways can lead to slightly different intensities among its resonant modes, but it will not change the frequencies at which the object vibrates or, in general, the relative degrees at which an object vibrates at those modes. Striking a coffee mug on the handle produces different intensities of the mug’s resonant modes than striking it on the rim does, but it does not change those frequencies.4

Philosophers have missed the relevance of stimulants like thwacks for understanding sounds. O’Callaghan, discussing Pasnau’s view of sounds, is a good example of this:

Full spectrum illumination (which daylight approximates) therefore has normative significance in revealing the colors of things because when reflected it carries information about how objects tend to interact with light across the entire spectrum….

No such normatively significant medium exists in the case of sounds. Neither air nor water nor helium optimally reveals the subtle vibrations of an object in the way that full spectrum light does for the reflectance characteristics of a surface. (O’Callaghan 2007, 54)

Because light requires no medium through which to propagate, and perhaps because of metaphors like being bathed in light, this is a reasonable, but misguided, way to relate sounds and colors. It confuses the medium (air) with the stimulant (a thwack). Objects are stimulated to vibrate by having energy imparted to them, and the normatively significant, ideal way to figure out what an object sounds like is to give it a dose of white noise: thwack it. White noise stimulates the object with equal energy across all frequencies and thus guarantees that all of the object’s resonant modes have a chance to vibrate at their characteristic levels. Stimulating an object with only one frequency gives a poor sense of its vibratory characteristics, just as stimulating a colored surface with a laser interferes with discerning the surface’s color. Where should one thwack it? Well, all over and at once, if possible, but ordinary thwacks with objects are fine substitutes for this rarely realized ideal, because objects vibrate quite similarly across kinds of stimulation. How hard should one thwack it? Hard enough, but remember: you break it, you bought it. An overly timid thwack will not cause the object to vibrate with audible intensity across the range of its resonant modes, so it poorly reveals the object’s sound just as sufficiently dim light poorly reveals an object’s color.

Because compression waves require a medium, there is another dimension constitutive of ideal audible conditions. Is it really true that water and helium cannot be excluded as ideal media? The obvious choice for an ideal medium is Earth’s atmosphere within a certain range of pressures, rather than, say, the surface of Titan in a methane rain. Even Earth’s atmosphere is problematic when there are high winds or when many other sounds are being made (see Benade 1976, 287). It helps to notice that the atmosphere makes a difference to color perception as well, even though light does not propagate within the air. Mountains look purple on a sufficiently hazy day.

In sum, the proposal is that an object’s sound is its natural frequencies of vibration in their ordinary proportions with respect to another. These dispositions are relatively stable across media — air, water, even the vacuum — across many kinds of stimulation, and they endure even when the object is unstimulated and thus does not vibrate. Objects that have distinct natural frequencies tend to sound different, and we can identify and recognize objects on the basis of how they sound.5 If you want to know how an object sounds, you need to

4. These natural frequencies are relatively stable across kinds of stimulation. It’s possible to thwack an object so as to make it vibrate, especially in the very short term, in an uncharacteristic manner (Benade 1976, 44) and, as we will see in Section 3, it’s possible to drive objects to vibrate in uncharacteristic ways as well.

5. Objects can sound indistinguishable even though they have different natural
objects’ sounds brought on by the uncharacteristic pressure waves they produce. The most obvious examples of this phenomenon are loudspeakers. What does a speaker sound like? If you are asking about the speaker’s sounds, the right answer is “Not much.” Test this by thwacking your speakers (be gentle!). At best, you get a dull, short-lived thud. Speakers respond equally poorly to stimulation across the range of audible frequencies. This fact about speakers makes it easy to drive them to make the sounds you want them to make. A speaker can vibrate as a guitar would vibrate even though it doesn’t have the sound of a guitar: an impressive piece of trompe l’oreille. One might object that we don’t ordinarily take speakers to be generators of illusions, but practice suggests otherwise. As speakers multiply from stereophonic to “surround”, we expect nothing less than the illusion not only that some object has a sound it does not really have but that we are hearing the sound of an object where there is in fact no object. Speakers sound like very little for the same reason that the screen in a movie theater is not painted bright red. Watching a film, we are as uninterested in the sounds of the speakers as we are in the silver of the screen. It’s the impressions that matter.

This distinction between having sounds and making sounds figures in two objections to the stable property view. There is little reason to deny that objects vibrate at natural frequencies, at certain relative strengths, when stimulated, since this has been a scientific staple since Helmholtz (1877). But these objections suggest it’s more plausible that we hear sounds made, and thus that sounds are vibratory properties or events, than it is that we hear stable dispositions of objects to vibrate. The first objection does this by pointing out that no similar distinction is drawn between having colors and making colors when stimulated by light (see Urmon 1969, 126). Colors really seem like “stable, lasting properties,” to use Pasnau’s expression, while sounds do not. The second objection is that objects typically make sounds other than the ones they have. So it is implausible that what we hear are stable dispositions to vibrate, and much more plausible that we hear the vibrations
that objects happen to have. The next paragraph addresses the first objection, while the remainder of this section and the section to follow address the second.

We seem happy to distinguish between the colors an object has and the colors it makes, even though we don’t talk that way about it. For example, CDs don’t seem to be rainbow-colored even though rainbow experiences are common when looking at CDs. Hylas notes that the colors on a pigeon’s neck seem rather unstable (Berkeley, Dialogues III). A red car can reflect the blue sky and green trees in such a way as to provoke experiences of green and blue even though the car is experienced as being red. Silvery things look silvery even though they generally evoke experiences of color by reflecting light from other objects. Weird lighting can make scenes seem a mix of oddly colored and oddly illuminated (see Hilbert 2005, §3.1). Rainbows evoke color experiences even though they are not experiences as of objects in any ordinary sense of the word. In short, there are many color experiences that are not experiences of objects’ colors or even experiences as of objects’ colors, and this does not by itself undermine the view that colors are dispositions to reflect light.

One might retort, and this is the second objection, that what makes sounds distinct from colors is the relative prevalence of one kind of experience over the other. There are color experiences that are not experiences as of objects’ colors — in the sense of being experiences as of stable properties of objects’ surfaces — but these seem exceptional even if they are not exactly rare. By contrast, the norm for audition is not experiences as of the sounds an object has. Most sound experiences seem like the transient rainbows adorning CDs’ surfaces — sounds merely made — and if most color experiences were like this no one would think to identify colors with stable surface properties. Urmson thought that the closest visual analogs to sounds were not colors but transient events like “glows, flashes, glares, beams of searchlights and headlamps, and halos” (1969, 124). If Urmson is right, we should not identify sounds with objects’ natural frequencies of vibration. In support of this worry, consider the following examples.

Quaking aspens are so named for the distinctive sound they make when the wind rustles their leaves. Do aspens have the sounds they make? Leaves are not elastic objects of the sort that have resonant modes of vibration, so according to the present account they have no sounds. Similar remarks apply to shaking folds of fabric and crinkling paper. Even objects that have natural frequencies of vibration are often identified by sounds that they do not have. Twigs snap underfoot and silverware clangs in the sink. Twigs and forks have resonant modes, but the snaps and clangs do not reflect them well. It might be that most of what we hear is not natural frequencies of elastic objects, so perhaps much of what we hear is not heard as the sounds of objects in the sense discussed above. This undermines an attempt to understand sounds as stable properties of objects.

This objection foregrounds a lacuna in the view that sounds are natural frequencies of vibration, but it does not defeat the proposal. All objects have characteristic reactions to being struck, regardless of whether they tend to vibrate well at natural frequencies. When struck, an object vibrates in an odd but characteristic fashion for a brief time before it settles into vibrating at its natural frequencies, if it has any. Like natural frequencies, these attack patterns are relatively stable dispositions of objects to vibrate when thwacked. Similarly, objects’ vibrations decay characteristically. Attacks and decays are stable dispositions of objects to vibrate when thwacked, just as natural frequencies of vibration are. Since attacks and decays are things that happen over time, any manifestation of such a disposition must evolve over time. That does not suggest that sounds are transient so much as it suggests that manifestations of certain dispositions are transient.

Attack patterns are quite important audible features of objects. Kenneth Berger (1963, and see Rossing et al. 2002, 140) recorded wind instruments and then removed the first half-second of each note

6. In an analogy that bears repeating, Urmson claims, “Philosophers have too often talked as though colours were like glows, emitted by a thing but condensed into a film on the thing’s surface. It is as if a glow were like a pan of gravy and a colour like the film on the bottom of the pan when one has allowed the water to boil away. But colours are not dehydrated glows” (1969, 126).

7. I owe this example to Alex Barnett.
recorded. This, in effect, removed the attack from the recording. Under these circumstances, it was almost impossible to identify the instrument making the note, even for members of a university marching band who were quite familiar with such instruments. The one exception to this seems to have been the oboe. Listeners’ difficulties under these circumstances are unsurprising, since wind instruments are rather similar to one another in their natural frequencies while their attacks differ significantly from one another.\(^8\)

In a similar fashion, though leaves and paper can have no characteristic modes of vibration, they do vibrate in characteristic, if transient, ways when stimulated. If thwacked, they will react. These reactions are determined by the objects’ material constitutions and structures, and they are good candidates for sounds that such objects have. Rather than understanding objects’ sounds merely in terms of their natural frequencies, it’s better also to include their dispositions to react transiently to thwacks. If objects’ colors went through a brief, odd transition right after one turned on the lights this would not undermine an account of color in terms of surface spectral reflectances, but it would suggest a similar amendment, especially if such transients helped to identify objects visually.

Where does this leave us with respect to the objection that most sound experiences are not experiences of objects’ sounds or even as of objects’ sounds? Including characteristic attack and decline patterns within the class of sounds suggests that sounds made but not possessed are exceptions and not the norm. Now and again one encounters objects making sounds they do not have — loudspeakers do this, for example — but this is not the ubiquitous experience one might think it is, unless one spends all day listening to the radio. But remember, someone who was always in the cinema would be in a poor environment for discerning the nature of color. Generally speaking, objects can be stimulated in ways that are far from ideal for revealing the sounds they have. This is unsurprising, and it does nothing to undermine the analogy between odd sound experiences and the odd color experiences, like the rainbows on a CD.

Even given this more expansive set of vibratory dispositions, it remains true that there is variability in the ways objects vibrate when they are thwacked. The next section defends the idea that sounds are stable dispositions of objects to vibrate by considering auditory constancy. Constancy phenomena clarify the sense in which sounds are experienced as being characteristic, stable features of objects, rather than as transient disturbances thereof.

4. Auditory constancy

It is an understatement to claim that not everyone thinks of colors as Alex Byrne and David Hilbert do, as dispositions to reflect light across the range of visible frequencies.\(^9\) There are many avenues along which one can object to such a view, but it pays to consider what motivates Byrne and Hilbert to accept it. Color experience does not make it manifest that colors are such properties any more than auditory experience makes it manifest that sounds are dispositions of objects to vibrate. An important motivation for such a view of color is the phenomenon of color constancy (Byrne and Hilbert 2003, 9; Hilbert 2005 and 1992, 364–367).

We see things under many different kinds of illumination. The spectral composition of daylight at dawn is rather different from light at noon, or the deliverances of myriad bulbs. If the illumination is particularly bad, in that it lacks power across large swaths of the visible spectrum, color vision can be quite unreliable. To a large extent, however, the ability to discern the colors of things remains intact across changes in illumination. “[I]n many circumstances the perceived color of a surface will be more closely correlated with the (stable) reflecting properties of that surface than it will with the (variable) spectral power

---

8. Part of what is surprising is that these students performed poorly in identifying unaltered playbacks of these instruments as well (Berger 1964, 1889).

9. See, for example, the open peer commentaries on their article (2003, 22–51). Scientists seem to be those least pleased with reflectance physicalism, though prominent philosophical detractors include C. L. Hardin (1993) and Mark Johnston (1992). Prominent supporters of the view — e.g., Dretske (1995) and Tye (2000) — are not hard to find, either.
distribution of the light reaching the eye from that surface” (Hilbert 1992, 364–365, and see Hilbert 2005). The green grass looks green in the orange light at dusk, but it looks like green grass illuminated oddly. These constancies suggest to Hilbert, and to Byrne, that “the function of color vision is to determine aspects of the reflecting properties of distal surfaces” (Hilbert 1992, 365, and see Byrne and Hilbert 2003, 9). It would be difficult to explain color constancy if the function of the visual system were to detect the spectral constitution of light reflected from objects, since constancy interferes with that.

Sound constancy of the right sort would suggest that the auditory system is structured so as to track features of objects that are stable across different circumstances. Not all auditory constancies favor the stable property view over its competitors, however. The ones that matter are those that suggest objects sound the same across changes in the way that they are stimulated to vibrate, and thus across some changes in the pressure waves that they produce in response to those stimuli. Such constancies suggest that the function of the auditory system is not detecting the vibrations an object happens to have at a given time but detecting features that are relatively stable across different kinds of vibration resulting from different kinds of stimulation.

What is a kind of stimulation, anyway? Quite generally, it is a way of imparting energy to an object at different intensities across a range of frequencies. A sinusoidal pressure wave has the lion’s share of its energy centered around its dominant frequency, for example. Such a stimulant can cause an object to vibrate, but it reveals the range of an object’s vibratory dispositions only poorly. Between white noise and a sinusoid is a large range of stimuli that are better or worse at revealing objects’ sounds. The question about constancy is whether, perceptually, objects seem to have the same sounds across decent ranges of stimulants. Do objects sound roughly the same whether it is white noise or something reasonably far from it that stimulates them? It turns out that they do.

A compelling example in this vein is that we are fairly good at identifying individuals, and more generally kinds of voices, on the basis of hearing people speak. Speakers’ vocal tracts vibrate in innumerable ways during the course of any conversation. All of us have some range of frequencies that we can make merely by changing the tension of our vocal cords. This vibration stimulates the different resonant modes of the nasal cavities, lungs, and mouth. What we hear is rather complex, but it includes the sound of the vocal cords as well as the sound of the rest of the person, even though the primary stimulus for the latter comes from the sound made by the former. We hear the vocal cords vibrate, and we hear the rest of the person vibrate. We are able to identify the speaker even though she changes the characteristics of the sounds she makes in order to speak. This is exactly the kind of constancy that tells in favor of the present approach to sounds: we identify the sound of a speaker across many kinds of stimulation provided by her vocal cords. All of these stimulants contain a broad range of frequencies, but none of them are white noise.

What exactly are the constancies involved in the above phenomena? Most sounds we hear are complexes of frequencies across the audible spectrum because most objects that vibrate do so naturally at many frequencies at once. Each of these frequencies at which an object vibrates are “partials” of its sound. Sounds we hear are by and large composed of many partials. Indeed, objects that vibrate at some frequency \( f \), but nothing lower than \( f \), also tend to vibrate at \( 2f, 3f, 4f \), and so on. The lowest, “fundamental,” frequency is usually accompanied by these integer-multiple “harmonics” among other partials. An object has a pitch when it vibrates at a fundamental and/or a number of its harmonics or, alternatively, when enough of its partials are integer multiples of some other partial.\(^\text{10}\)

\(^{10}\) The vocal cords have different sounds, on the current view, depending on how taut they are. Generally, reorganizing the structural features of objects can change the sounds that such objects have.

\(^{11}\) This terminology follows Benade (1976, 63). Objects can be perceived to have the same pitch even though they differ in the frequencies at which they vibrate. Famously, removing the fundamental frequency from a stimulus, while leaving the other harmonics intact, preserves the perceived pitch of the stimulus. (See, e.g., Rossing et al. 2002, 126.)
The way an object sounds is not determined solely by its pitch, if it has any, but also by its timbre. Timbre is not a well-defined category within acoustics research, since it seems to include anything that is not pitch or loudness (Bregman 1990, 92–94). Timbre is determined by the partials at which an object vibrates, the relative intensities at which it vibrates at those partials, and any overall trends in the intensities of the partials. For example, a clarinet and a piano can make notes that have the same pitch, but the instruments sound different in part because of the relative intensities of the fundamental’s harmonics. Clarinets and other woodwinds, for example, have much more energy at the odd-numbered harmonics than at the even-numbered ones.12 Also, for most objects the intensities at different partials tend to decrease as frequency increases. So, even though the clarinet has relative peaks in intensity at the odd harmonics, the overall trend of these intensities is downward. This overall trend — the spectral slope — is modulated by a particular envelope shape, which is a general rising and falling pattern among the partials as frequency increases.

Much of the foregoing holds for human voices. The pitch of one’s voice changes over the course of any conversation. Timbre changes too, as one vocalizes different phonemes. We are very good at recognizing phonemes regardless of who is speaking (Gagne and Zureck 1988), but we are also good at identifying speakers regardless of the phonemes they vocalize or the pitch at which they vocalize them. It turns out that the rate at which partials decline in amplitude as frequency increases — the spectral slope of a signal — is quite diagnostic for identifying kinds of voices. Citing Monsen and Engebreston (1977), Xiaofeng Li and Richard Pastore point out “[A] breathy voice tends to have a steeper spectral slope than a modal voice, whereas a creaky voice has a shallower spectral slope than a modal voice…. In addition, spectral slope differs between speech sounds of different genders

12. These instruments can be heard to have the same pitch despite sounding different. Humans have decent “pitch constancy” (Warrier and Zatorre 2002), which is enhanced when sounds are heard within melodic contexts. This kind of constancy is not diagnostic for whether the present approach, O’Callaghan’s, or Pasnau’s is best.

with female voices typically exhibiting steeper spectral slopes than male voices” (1995, p. 1957). Part of how we identify speakers across the range of noises they make is in terms of the spectral slopes of their voices. Li and Pastore show that we are quite good at identifying sounds in terms of their spectral slopes, independent of other changes in them. The particular phonemes one produces depend more on the spectral envelope pattern than on the spectral slope.13 Both envelope pattern and slope are fairly constant across changes in the fundamental frequency produced by the vocal cords. All of this suggests that we are good at identifying both individuals and those configurations of their faces that make individual phonemes across kinds of stimulation, since the stimulant in these cases is the vibration of the vocal cords. There is no obvious reason the auditory system should exhibit constancy for something like spectral slope or envelope pattern, unless the auditory system has the function of identifying stable properties of objects.

Speech perception is a very special aspect of audition, akin to the way in which face perception seems special visually. The fact that certain things are true of face perception does not mean that they characterize vision generally because there is evidence that faces are processed in a special way, perhaps with dedicated hardware (Kanwisher et al. 1997). Many think the same is true of speech. Li and Pastore show that this feature of audition that allows one to distinguish speakers and kinds of voices is also present for synthesized stimuli not encountered in nature. There is robust spectral slope constancy for synthesized sounds.14 In similar work using synthesized tones, Jean-Pierre Gagne and P.M. Zureck (1988, esp. 2298) found that we are actually good at

13. There is no simple account of how we recognize phonemes across different speakers and different contexts of utterance (Magnuson and Nussbaum 2007). The details of such an account might tell against the view proposed here, but so far nothing seems to contradict it.

14. O’Callaghan suggests that constancy for some aspects of timbre favors his event view over the wave view (2007, 89). But these constancies do not recommend the event view over the stable property view. At best, they show that either should be favored over a wave view.
detecting changes in the properties of objects that are stimulated differently relatively independently of the character of those stimuli (also see Bregman 1990, 98–103). Handel (1989) and Ballas (1993) concur that the results for speech sounds apply more generally.

Something similar seems to happen with many instruments, including guitars, violins, and pianos. The individual notes played are determined by the length, width, and material of the stimulated string, but the vibrations from the string excite the body of the instrument to vibrate. The main stimuli for the instrument are the many vibrations produced by the strings. We distinguish the sound of a guitar from the individual note it happens to play. Those with keen ears can even distinguish individual instruments from one another. In a related vein, Daniel Freed showed that we can identify the hardness of a mallet that strikes a pan across different kinds of pans and we can identify pans across kinds of mallets. “Evidently, subjects are able to focus on the mallet-related aspects of the stimuli and ignore the pan-related aspects (such as pitch)” (Freed 1990, 319).

These results reflect precisely the kind of auditory constancy one would expect if sounds are stable dispositions of objects to vibrate in response to being thwacked. They suggest that auditory perception concerns not how some object happens to be vibrating at any given moment but rather how that object is disposed to vibrate across modes of stimulation.

None of these results are inconsistent with Pasnau’s vibration view, O’Callaghan’s event view, or even a wave view of sounds, but they favor the stable property view. In a recent article, Pasnau suggests that we should conceive of both colors and sounds as events rather than as stable properties—perhaps colors are dehydrated glows after all—but he acknowledges that the event view of color is threatened by color constancy (forthcoming, 55). Constancy is consistent with his view of color, but it is more readily handled by a stable property view. More generally, the best explanation of the auditory and visual constancies discussed here and by Byrne and Hilbert is that perceptual systems latch onto properties of objects stable across modes of stimulation.

5. Hearing events

A potential problem for the present account is the fact that sounds often seem to inform us about events. “[A] given sound provides information about an interaction of materials at a location in an environment” (Gaver 1993, 6: Gaver’s italics). This suggests that the function of hearing is at least in part informing perceivers about happenings rather than about objects’ stable vibratory predilections. If so, this might favor an account of sounds as events. Albert Bregman, O’Callaghan’s principal source of empirical support for his view, claims that “[a]coustic information … tells us about physical ‘happenings’” (Bregman 1990, 10). We hear hands clap (Repp 1987), bottles break, and balls bounce (Warren and Verbrugge 1984). Indeed, “it is arguably among the primary functions of auditory perception to informs us about the temporal characteristics, including the durations, of happenings in our environment.” (O’Callaghan 2007, 45). How, then, could sounds be stable properties of objects?

One way of approaching this question is to ask whether and under what circumstances we manage to see events, how this relates to color perception, and how that carries over to the auditory case. Turn on the lights and you notice that the room has become illuminated, in part because you can now discern things’ colors. Change the spectral composition, brightness, orientation, or spatial extent of illumination and you notice these events too. These events do not seem to affect the
colors of things, but they crucially involve perception of things’ colors. Spotlights move quickly across the room. Phenomenologically, perception of their motion is tied to the revealing of things’ shapes and colors as the lights move across the otherwise dark space. In an extreme, rave-inspired version of this scenario, lights of many colors twirl, pulse, blink, brighten, and dim, making everything seem like it’s something that’s happening. We see many events by seeing things’ colors.

Things are similar with sounds, or so the current view would suggest. We hear events by hearing the sounds of things. The mallet strikes the pan. Both objects are heard and one hears the event of the mallet striking the pan, somehow, by hearing both objects sound off at the same time. As Freed (1990) showed, we hear the pan despite some differences in its actual vibrations that result from differences in the mallets used. Exactly how the perception of an event per se proceeds from the perception of things’ sounds is as (un)clear as how the perception of events relates to the perception of things’ colors. James Ballas (1993, 257), for example, found that listeners were equally good at identifying sounds in terms of an action involving certain objects as they were at identifying the objects involved in the action. This is not conclusive one way or another, but it lends plausibility to the thought that we hear events by hearing stable properties of objects. Combined with the facts about auditory constancy discussed above, the case becomes even more plausible.

Even accounts of sounds as events need to explain how we hear the events we ordinarily understand ourselves to hear. For O’Callaghan, the sound is the disturbing of a medium by an object, but a collision between objects is not the disturbing of a medium by either of those objects. The events are related, even though they are not identical. So O’Callaghan needs to show how the perception of interesting events happens on the basis of sound perception. Similar remarks apply to proponents of the wave view, like Sorensen, or proponents of other event views. So, the fact that we hear events does not obviously favor one account of sounds over any of the others.

One thing that clearly distinguishes perception of color from perception of sound is the way in which the stimuli work in each case. Light is generally diffuse and thus illuminates most objects in an environment indifferently. In addition, light is long-lasting in that illuminants tend to keep on shining. Thwacks, by contrast, are focused and brief. We rarely hear most things in a room at once, and we rarely hear them for long in any case. The visual analog of this would be if we were only able to see via short bursts of rather focused light. In that case, the relation between seeing these objects like colors and seeing events would seem more pressing, since each occasion on which a color was seen would also be identifiable as an event in which the object was illuminated. This would not undermine an account of color properties as stable dispositions to reflect light. Indeed, what would matter most in such a situation are the constancies that characterize such visual perception. Color constancy being what it is, there would be a strong case in favor of identifying colors as stable properties of objects, and the same seems true in the auditory case.

Finally, this discussion of event perception is relevant to the practice of counting sounds. Strike the table twice and you hear two sounds, right? The stable property view denies that claim. You hear one sound twice. This strikes many as counterintuitive, but it helps to notice that everyday practice is not uniformly aligned with the view that sounds can be counted in the manner suggested. Many locutions — “Let me hear that sound again,” “I could listen to the sound of that bell all day” — are better aligned with the stable property view. They suggest that what we count are the hearings of sounds or the occasions on which a sound is made audible — occasions when an object vibrates in a medium — rather than the sounds themselves. In contrast to colors, the occasions on which a sound becomes audible and during which we hear sounds are focused, brief, and much more compellingly countable. So it is understandable that we talk about sounds...
as things that can be counted, but this practice does not constitute a worked-out philosophical account of sound. Indeed, practice pushes us in a number of directions at once. The considerations canvassed so far lend support to the stable property view that the habit of counting sounds does not undermine.

6. Loudness

The stable property view suggests that sounds are stable dispositions of objects to vibrate. Auditory constancy favors such a view because the phenomenon suggests that we track stable features of objects, abstracting over peculiarities of how an object vibrates now and then. But the stable property view also suggests that the content of auditory experience is complex in a particularly interesting way. Specifically, nothing has so far been said about loudness. Loudness clearly plays a role in auditory experience, but it has little to do with the frequencies at which objects vibrate. Only relative intensities play any role there. Loudness is not a feature of sounds for the stable property theorist, which suggests that there is more to auditory experience than sound. Or, if one finds that way of talking unappealing, since ‘sounds’ should cover the full range of what is heard, one could say that there is more to sound than stable dispositions of objects to vibrate.

How should one think about loudness? Turning the volume to 11 does not change the sounds heard; it changes how loudly they are heard. Loudness is an aspect of the circumstances under which sounds are heard. Each occasion on which a sound becomes audible is one on which some stimulus causes an object to vibrate. The more energy the stimulus has, the louder things get. Loudness is the intensity of such stimulating events. While it is common to talk of sounds as having pitch, timbre, and loudness, on the current view sounds have pitch and timbre, but loudness is not a feature of the sound. It’s a feature of the thwack that reveals the sound.

As mentioned in Section 2, if an object is not thwacked hard enough, it will not vibrate audibly at all of its natural modes. We are in a poor position to discern things’ sounds when things are not stimulated with sufficient intensity. Loudness perception provides a sense of how good the conditions are for hearing something’s sound. In addition, loudness perception figures in the perception of events in that loudness is a feature of events that make sounds audible.

There is some precedent for this treatment of loudness in a recent discussion of color constancy. Hilbert suggests that it is wrong to think of surfaces’ visual appearances solely in terms of hue, saturation, and brightness. Indeed, “Those color scientists interested in the details of color appearance have been aware for some time that three dimensions are not enough to fully characterize color appearance” (Hilbert 2005). He cites, in particular, M. D. Fairchild’s (1998) textbook on color appearance models as suggesting five dimensions instead of three. Why broaden things out in this manner? Because, as mentioned in Section 4, the green grass at dusk looks like green grass illuminated oddly. The grass at noon on a sunny day looks like grass illuminated brightly, while on a cloudy day it looks more softly illuminated. These are all aspects of the object’s appearance, and some of them concern the illuminant. Visual experience reveals aspects of the illuminants that allow one to see color in the first place. Among other things, perception of illuminant intensity gives us a sense of how good conditions are for seeing objects’ colors. Hilbert concludes that “it is time to complicate our picture of the content of visual experience and ... this more complicated content allows us to provide a better understanding of visual phenomenology” (2005). The same holds true for auditory experience.

Thinking about sounds as stable properties forces one to think hard about auditory phenomenology. Section 2 pointed out that philosophers have completely ignored the role that stimuli like thwacks play in auditory perception. No wonder that no role analogous to that of the illuminant in vision has been suggested for audition. Once con-
stancy phenomena are given their due and sounds are thought of as stable properties, it is actually quite natural to regard loudness as akin to the intensity of the illuminant in color vision. In fact, if stimulants are relevant in the way suggested herein, and sounds are analogous to colors, one should expect aspects of stimulus intensity to manifest themselves in experiences of sound. Furthermore, it is phenomenologically compelling to separate loudness perception from the perception of pitch and timbre. Loudness seems more a contingent aspect of sounds than pitch or timbre does. Change the latter and you change the sound, but change the former and you just change how loud the sound is. These phenomenological claims are far from conclusive, but they are worth mentioning lest one think it is obvious that loudness is just an aspect of sounds no different from pitch or timbre.

7. Audible qualities

One might object at this point that the deepest problems for the stable property view remain untouched. Pitch, timbre, and loudness are correlated with aspects of the frequencies and intensities with which an object vibrates or is disposed to vibrate, but that does not mean we can just identify these audible qualities with their physical counterparts.

Reflectance physicalism is controversial in part because, its detractors claim, it does a poor job of accounting for the manifest image connected with colors. Colors do not seem like reflectances, and many of the relations that colors bear to one another are not straightforwardly reflected in relations between reflectances (Hardin 1988, e.g.). Similarly, one could argue, pitch is related only in a very complex way to frequency. Perception of a rather unified-sounding pitch is the perception of quite a broad range of frequencies at different intensities. There is nothing in auditory experience that would lead one to suspect this is so, which calls any identification of pitch with complexes of frequencies into question. Similar remarks apply to timbre and even to loudness, which is not as straightforwardly related to the amplitude of compression waves, or thwacks, as one might have thought (Rossing et al. 2002, 108–112). Some auditory scientists would doubtless claim that audible qualities are not to be found in the world. Instead, they are properties somehow constructed or invented by the brain on the basis of its encounters with pressure waves. This is exactly how many color scientists feel about color (e.g., Palmer 1999).

It’s easy to agree that these are issues with which any complete and convincing account of sounds must deal, but for now they are beside the point. The conflict in this paper concerns those accounts that identify sounds as transient properties of objects or as events and the present one that identifies sounds as stable dispositions to vibrate. Pasnau and O’Callaghan are both physicalists about sounds, in the sense that they think sound, pitch, timbre, and loudness are mind-independent entities that admit of straightforward scientific explanation. In fact, O’Callaghan provides a rather compelling case for identifying pitch, timbre, and loudness with physical properties (2007, Ch 6). His defense is largely independent of his account of sounds as events and it is informed by reflectance physicalist accounts of color. Given the rather tight analogy that this paper draws between colors and sounds, one would expect a rather close relation between the strategies for defending reflectance physicalism about color and defending the present view.

8. Conclusion

It is intuitive to think of sounds as transitory entities. Either they are properties that objects have while we hear them, or they are events: time-like particulars. Indeed, O’Callaghan (2007, §3.7) thinks that one of the chief desiderata a theory of sounds must satisfy is that it must render perceptual claims about sounds’ durations veridical. Sounds seem to last mere moments, so the best account of sounds should suggest that they do in fact last mere moments. Upon reflection, however, it far from obvious that things seem this way. We generally hear sounds for mere moments, but much suggests that the sounds themselves endure. Few of the phenomena that Pasnau or O’Callaghan mention are inconsistent with the claim that sounds are stable properties, and the facts of auditory constancy strongly favor a stable property view. It’s
certainly true that sounds inform us about events and even about the
durations of events, but we are similarly informed about events by
the perception of colors over time. In any case, even event views of
sounds put perception of the events of interest at one remove from the
perception of sounds.

Thinking of sounds as stable properties depends on exploring the
points of analogy between auditory phenomena and visual ones. This
proposal is not slave to a visuocentric view of things, though it does
take some pointers from the much better-developed philosophical lit-
erature concerning colors. Perhaps blindness (or is it deafness?) to the
salient points of analogy between colors and sounds is one of the chief
reasons this kind of approach has not even been considered, let alone
favored. Indeed, no worked-out account of sounds has ever drawn a
tight analogy between sounds and colors. The hope is that by taking
the analogy between vision and audition seriously, one finds a nu-
anced account of sounds, with the added bonus of a unified picture of
the two modalities.19

References

Ballas, J. A. 1993. Common factors in the identification of an assort-
ment of brief everyday sounds. Journal of Experimental Psychology:

Dover.

Berger, K. W. 1963. Some factors in the recognition of timbre. Journal of
the Acoustical Society of America 36: 1888–1891.

19. This paper benefited substantially from the generous help of others. Casey
O’Callaghan, Roy Sorensen, and anonymous referees for this journal provid-
ed very helpful comments on earlier drafts. Carl Matheson commented on a
version of the paper at the 2008 Canadian Philosophical Association meeting.
My colleagues at Dartmouth and an audience at Carleton University — espe-
cially Larry Crocker and Heidi Maibom — provided helpfully skeptical feed-
back. Alex Barnett was good enough to share his knowledge of the physics
of sounds and sound waves. Finally, a Dartmouth Junior Faculty Fellowship
provided me with research time to complete work on this project.

Indianapolis: Hackett.

Press.


Wesley.

Freed, D. J. 1990. Auditory correlates of perceived mallet hardness for
a set of recorded percussive sound events. Journal of the Acoustical
Society of America 87: 311–322.

Gagne, J., and P. M. Zurek. 1988. Resonance-frequency discrimina-

Gaver, W. 1986. Auditory icons. Human-Computer Interaction 2:
167–177.

——. 1993. What in the world do we hear? An ecological approach to

Handel, S. 1989. Listening: An Introduction to the Perception of Auditory
Events. Cambridge, MA: MIT.


London: Longmans, Green.


Topics 33: 141–158.

Press.

68(3): 221–263.


——. Forthcoming. The event of color. *Philosophical Studies*.


Ungerleider, L., and M. Mishkin. 1982. Two cortical visual systems. In...