Audition and Composite Sensory Individuals

**Abstract**: What are the sensory individuals of audition? What are the entities our auditory system attributes properties to? We examine various proposals about the nature of the sensory individuals of audition, and show that while each can account for some aspects of auditory perception, each also faces certain difficulties. We then put forward a new conception of sensory individuals according to which auditory sensory individuals are *composite* *individuals*. A feature shared by all existing accounts of sounds and sources is that they postulate sensory individuals that are non-composite. They identify the sensory individual of sound hearing or source hearing with one type of entity in the environment, be they sound waves, vibrations, or interactions. We question this assumption and argue that our perceptual systems represent two or more aspects of the environment as a single sensory individual. Finally, we show that taking auditory individuals to be composite sensory individuals allows for an account of audition that is less problematic than its existing alternatives.

## Introduction: The Sensory Individuals of Audition

Perception is often thought to be, at root, a matter of (conscious or unconscious) *property attribution*, with properties like colour, shape and texture, being attributed visually, and properties like pitch, timbre, and loudness, being attributed auditorily. Our focus here is not on which properties are auditorily attributed (a longstanding and high-profile debate in philosophy of perception, see, for example, Siegel 2006, Nanay 2011, and, in the case of audition di Bona 2017), but rather on what those properties are attributed to. That is, on the *sensory individuals* of auditory perception.

The sensory individuals of vision are commonly thought to be ordinary material objects (for versions of this view see Matthen 2005, pp. 277–282; Nanay 2013, p. 51; Cohen 2004; for an opposing view, see Clark 2000, ch. 2; 2004), but when it comes to auditory individuals, there is far less consensus. We might think that this is surprising, and that it is obvious what the sensory individuals of audition are. We hear *sounds*. Such a straightforward answer, however, faces various difficulties.

For a start, there is disagreement as to what sounds are. Which type of environmental feature is picked out and represented? The most obvious, common sense, answer to this question would be ‘sound waves’ but such a view has been subject to various criticisms (REF) and various alternatives have been put forward, identifying sounds with vibrations (Casati and Dokic), or disturbances of the air (or some other medium) caused by these vibrations (O’Callaghan). Opinions differ as to what *type* of feature sounds are. Are they *particulars* which are attributed pitch, timbre, and loudness, or are they *properties* which are themselves attributed to particulars?[[1]](#footnote-1) Rather than entering into these debates here, we will assume in what follows that hearing a sound is a matter of representing some type of particular and focus instead on the sensory individuals of *source hearing*. In contemporary philosophical work on audition it is very common to say that we not only hear sounds, but their source events: the collisions, rollings, and scrapings, which cause objects to vibrate and produce sound waves in the surrounding air. It is not obvious how such a claim can be understood in terms of sensory individuals. First, what aspect of the environment, specifically, do we represent and attribute properties when we hear sources? Second, how does the hearing of sources connect to the hearing of sounds? That is, how does the representation of whatever aspect of the environment we identify with sounds, relate to the representation of whatever aspect of the environment we identify with sources?

In what follows, we shall examine various proposals about the nature of the sensory individuals of source hearing, and show that while each can account for some aspects of auditory perception, each also faces certain difficulties. We then put forward a new conception of sensory individuals according to which auditory sensory individuals are *composite* *individuals*. A feature shared by all existing accounts of sounds and sources is that they postulate sensory individuals that are non-composite. They identify the sensory individual of sound hearing or source hearing with one type of entity in the environment, be they sound waves, vibrations, or interactions. We question this assumption and argue that our perceptual systems represent two or more aspects of the environment as a single sensory individual. Finally, we show that taking auditory individuals to be composite sensory individuals allows for an account of audition that is less problematic than its existing alternatives.

## 1. Sounds and Sources

### 1.1 Sources as Interaction Events or Vibration Events

Why think that we hear source events as well as sounds? First of all, source events stand in straightforward causal relations to auditory experiences. Second, we very often become aware from auditory experience of what sort of source event is taking place (e.g. a wooden window frame rattling, a glass shattering), immediately and without conscious inference. Third, it is plausible to suppose that the function of the auditory system, like all perceptual systems, is to provide us with useful information about material objects and their activities. We have an *interest* in what objects are doing what, rather than in how objects are causing the air around them to move.

While none of these reasons shows absolutely that sources are heard, they do at least show that accounts of auditory perception which allow for source hearing are desirable, and we shall assume this in what follows. We can think of the claim that we hear sources as a claim about sensory individuals. What aspect of the environment do we represent and attribute properties to when we hear sources?

Assuming that sounds are particulars, there are at least two versions of this view. According to the first approach, source hearing is a case of representing objects interacting or moving against one another. Consider, for example, the following remarks from O’Callaghan:

By sources I mean things like floorboards, electronics, collisions, jiggled keys and speeches. It is plausible that sound sources...are audible. Spoken language permits saying you hear floorboards, electronics, collisions, jiggled keys and speeches rather than just their sounds. Audition prompts you to form thoughts about material bodies and happenings. Certain aspects of auditory experience, such as the way things appear grouped auditorily, are attuned to features of material objects and events rather than merely to features of sounds. (O’Callaghan 2011, p. 376)

There are many different possible ways in which objects can interact in such a way as to produce sound waves (and thereby elicit auditory experiences), including collisions, rollings, scrapings or smashings. For simplicity, we will just use the label ‘interaction events’. Although there are disagreements as to exactly how the details should be fleshed out,[[2]](#footnote-2) thinking of source hearing as representing interaction events can be considered the mainstream view (see, for example Kulvicki; Leddington; Young; Nudds).

However, a slightly different conception of source hearing can be found in Casati et al.’s “ockhamization” of auditory perception (2013). They say:

#### The sound is not a proper part of a distinct event that is its source; it is identical with what we called the event source. The collision you hear is the sound you hear; the vibration of the stuff the tuning fork is made of is the sound. There is no difference between hearing the sound and hearing the event source, because there is no difference between the sound and the event source. (2013, p. 2)

We can see that the authors are sympathetic to the idea that to hear a source is to hear an interaction event. Indeed, they say that interaction events that we hear, such as collisions, just *are* the sounds that we hear. On the other hand, they also seem to identify sounds with another type of event: “the vibrations of the stuff tuning fork is made of”. That is, an event that occurs immediately after the interaction event. If both the collision and the vibration of the stuff are “the sound” then Casati et al. could be interpreted as saying that source events, as well as being interactions, are vibration events.[[3]](#footnote-3) However vibrations would seem to be a quite different type of event from interactions. Consider a visual analogy: I strike a tuning fork behind you and you turn to see the object still quivering from the impact. While you can see this *aftereffect* of the collision perfectly well, you did not see the impact itself. Regardless of whether Casati et al. really do think vibrations of stuff are sources instead of, or as well as, the interactions that produce them, these considerations do provide us with an alternative sensory individual for source hearing.

We might think that this alternative is less mainstream than the interaction view for good reason. Consider the motivations mentioned above for why an account of source hearing is desirable: sources stand in straightforward causal relations to auditory experiences, we become immediately aware of them through auditory experience, and we have an interest in them. While vibrations stand in straightforward causal relations to auditory experiences, we do not hear *how* objects are vibrating (for example, we do not become aware of the frequencies at an object is vibrating), and it is not as if we have any interest in the vibrations of objects *per se*.

On the other hand, identifying source events with vibratory events does have one advantage over thinking of sources as interaction events. Consider a very common –perhaps the most common– type of source event: collisions. When a tuning fork is struck, the auditory experience you enjoy begins only when the fork begins to vibrate, and only lasts as long as these vibrations continue. This is because –obviously– sound waves only propagate through the air and reach the eardrums while the tuning fork is vibrating. By the time the fork begins to vibrate and auditory experience begins, however, the interaction event is already over: the mallet has already *completed* its journey through the air to meet the fork. A reason to favour identifying vibrations with sources over the interaction events which they produce, then, is that vibration events are more *temporally aligned* with our auditory experience than interaction events.[[4]](#footnote-4)

### 1.2 Hearing Sources through Sounds

Whereas Casati et al argue that sound hearing and source hearing are one and the same, those who think that we hear source events, but that sounds themselves are sound waves, or medium disturbances, appear committed to the claim that audition involves the representation of *two* sensory individuals: sound and source. If this is correct, they are required to explain how the representation of one is connected to the representation of the other. As O’Callaghan puts it:

If you hear a source such as the jiggling of keys, then you may hear the jiggling keys to be the source of the jingling sound. When you do, you hear the sound to stand in some relation to its source. What relation? (O’Callaghan 2011, p. 2)

Nudds has argued that we attribute sounds with the property of *having been caused by* a source of a particular sort. He goes on to suggest that representing this sort of property allows for the representation of sources themselves:

#### When we experience a sound as having been produced by a source, our experience represents it as having that non-intrinsic property. Therefore, our auditory experience represents sounds and the sources of sounds and it represents sources as the sources of sounds by representing sounds as having a non-intrinsic property – the property of having been produced by a source of a certain kind (Nudds 2010b, p. 18)[[5]](#footnote-5)

However, even if we grant that sounds are heard to bear these types of properties, we might still doubt that this amounts to the sources of sounds themselves being *represented* in auditory experience. This is not simply the objection that perceiving the product of some cause is often not sufficient to see the cause itself (seeing smoke does not allow for the seeing of the fire; seeing a footprint does not allow for the seeing of the foot) (Dretske 1981, O’Callaghan 2011); rather, even if the result of some cause were perceived as bearing the non-intrinsic property of having been produced by that cause this would not amount to the perceptual representation of that cause. Imagine, for example, that apples from the south of Italy had the perceivable property of coming from the south of Italy, seeing apples to bear this property would give you good reason to think that they came from the south of Italy, but nothing here amounts to the perceptual representation of the south of Italy. In other words, regardless of the merits of Nudds’ proposal, we might prefer to find an account on which sources are auditorily represented in their own right.

O’Callaghan argues for an alternative *mereological* account of source hearing, on which we hear sounds as parts of sources and, through hearing these parts, *mediately* perceive sources (**2011**) . That is, hearing a sound allows for the hearing of a source in a manner somewhat similar to how we might think that seeing the front facing parts of an object allows for the perception of the *whole* object (e.g. Clarke 1965; **Bermudez REF**). Put in terms of sensory individuals we can say that, on this approach, we represent two features of our environment, sound and source, and the representation of the latter is secondary to, or depends on the representation of the former.

However, one worry we might have here is that seeing part of an object as a part would seem to involve seeing it as connected to other parts of the object, which could themselves potentially be brought into view: if you are looking at a table in front of you, you might see the tabletop and two of the table legs, and you see them as connected to other, occluded parts of the object, such as the underside of the table top, and the other two legs. If you changed your viewing angle, these parts could be brought into view (Nanay 2010, 2018). But audition seems quite different: if we hear the sound that the rattling window makes, it is not as if we can move (or do anything else) so that other parts of the source (such as the window frame or the collision event) are brought within earshot. If this is not possible, what reason do we have to think that we hear these sounds as parts connected to other –unheard– parts of sources, and therefore, what reason do we have to think that we are mediately perceiving the source after all? O’Callaghan responds to this type of worry by appealing to the possibility that auditory experience can be influenced by experiences in other perceptual modalities. He says:

the relevant possibilities for perception are not restricted to one sense. The events you hear are perceptible other than through audition, and this shapes the perceptual experience of hearing them. Hearing as of something that is or could be seen or touched makes a difference to the phenomenology of auditory experience...Hearing as of something that has visible or tactile features differs phenomenologically from hearing as of something without such features. Thus you might hear as of events that have inaudible parts. The audibility of an individual does not require that each of its parts is audible. (2011, p. 400).

O’Callaghan goes on to develop two ways in which experience in other modalities “shapes” auditory experience so that it allows for mereological source hearing. First, if we have visual or tactile access to the source event which is currently making the sound we hear, this could lead to a multimodal perceptual experience on which an individual is represented as bearing both visual and auditory features. Second, if we can only hear a source, but have had “a long history of both seeing and hearing” it, we “perceptually experience it to be something that has unseen but nonetheless visible features” (ibid. p. 401). While experiences in one modality are almost certainly influenced by present or past experiences in another, we might think that this type of response has an unappealing consequence: it limits the ability to hear sources to those that are able to have experiences in other perceptual modalities. Imagine a person who can hear perfectly well, but was born blind and with complete hypothesia (no sense of touch). As their auditory experiences are not, and would never have been, shaped by visual or tactile experience, it seems as if this would preclude them from being able to hear sounds as parts of sources, and so also preclude them from ever hearing sources. We might think that this consequence is too big a bullet to bite. Would a person in this position have such radically impoverished auditory experiences compared to ordinary perceivers? This seems especially pressing if we accept that the *function* of auditory experiences is to represent sources.

A second worry is that explaining the relation between hearing sources and hearing sounds in a way that does not appeal to perceived causation gives up too much. O’Callaghan makes clear that he is not denying that sounds are effects of sources, but is sceptical that they are heard as such (p. 401). But is this approach too revisionary? If we accept that we hear sounds and sources, and hear them as related, it is at least tempting to say that the relation we hear between the two is causal. Of course we should be wary of putting too much weight on phenomenological seemings, but, all else considered, an account which retains the idea that we hear sounds as caused by sources seems preferable to one, such as the mereological view, which does not.

We have outlined two issues about the auditory individuals of source hearing. First, there is the question of whether hearing sources consists in auditorily representing interaction events or vibration events. Second, there is the question as to the nature of the audible relation between sounds and sources.

## 2. Composite Individuals

### Composite Individuals

We can now see the difficulties that current accounts of source hearing face. Implicit in all of these views is the idea that the sensory individuals of audition match up with *one* type of environmental feature: source hearing is the representation of vibrations *or* of interaction events; we hear sources through sounds either by representing sounds to have a ‘having been caused by’ property, or by hearing sources *via* the hearing of sounds.

We suggest that if this assumption is given up, if we allow that two or more different types of environmental features can be represented as a single *composite* sensory individual, these problems can be avoided. Different theories of sounds and sources can be united in such a way as to retain their advantages and eliminate their disadvantages. In this section we shall outline how composite sensory individuals should be understood, in the next we shall show how they can be applied to the issues concerning sound and source hearing that have been introduced in the previous section.

Some particulars are not ontological simples – they are composed of different parts. In fact, most particulars are – maybe even all of them. A table is composed of the tabletop and the legs of the table. But the tabletop itself is composed of atoms, subatomic particles, etc. This is an ontological distinction between ontological simples and their mereological sums.

But we can ask a related question about how particulars are represented perceptually. When you see a table, what you see is a composite particular – it is composed of various parts. Is the sensory individual of this perceptual episode a composite sensory individual? It is, as long as your perceptual system represents the table as being composed of different parts. A sensory individual is a composite sensory individual if and only if the perceptual system represents it as being composed of different parts. Remember that a sensory individual is part of the perceptual content: it is what properties are perceptually attributed to. If properties are perceptually attributed to a composite entity, this constitutes a composite sensory individual. Lots of things are composite. But only if we perceptually represent them as composite can we talk about a composite sensory individual.

More needs to be said about what it means to represent something as composite. Suppose you represent particular X and X has some proper parts A and B. If representing X entails representing A and it also entails representing B, then X is represented as composite. Not all representations of composite particulars represent them as composite. If X, a composite particular, has some proper parts A and B, but representing X does not entail representing A or B, then while X is a composite particular, it is not represented as composite. To go back to the table example, a table is a composite particular, but we can represent a table without representing it as composite: if you tell me that you bought a table, I can form a belief that you bought a table without having any kind of representation about the parts of the table. Representing a table does not entail representing parts of the table.

The last paragraph was about representation in general: we can represent a composite particular without representing it as composite. But how about perceptual representations? Can we perceptually represent a composite particular without representing it as composite? While a case could be made that one of the important differences between perceptual and non-perceptual representations is that the former entails representing proper parts (if there are any), whereas the latter does not, we will not defend this claim here.

But while representing a table does not entail representing parts of the table, *perceptually* representing a table does seem to entail representing parts of the table. If you see a table, you represent both the tabletop and the legs. Maybe the legs are occluded, but you still represent them. If you did not represent the legs, you would not perceptually represent the table (but something else – the tabletop). So at least in many (but maybe even all) cases of perceptual representation, perceptually representing something entails representing its proper parts. Quick note about the scope of this claim: Does it entail representing all of its proper parts? Presumably not, but it entails representing at least some of its proper parts.

Given that, as we have seen, most (and maybe all) particulars are composite and in most (and maybe all) cases of perceptual representations, perceptually representing a particular entails representing its proper parts, it follows that most sensory individuals are composite. Here is a way of keeping composite and non-composite sensory individuals apart. If your perceptual state represents a composite sensory individual, you can attend to its proper parts. When you see the table, you can perceptually attend to the tabletop or the legs. Or some part of the tabletop. All these would amount to attending to parts of the composite sensory individual. If you could not attend to parts of a sensory individual, this would mean that either it does not have proper parts, or if it does, its parts are not represented. Either way, the sensory individual would fail to qualify as composite.

The question we now turn to is the ways in which a sensory individual can be composite. If a particular has proper parts and these parts are represented, we can ask what the relation is between these parts. We will talk about three options: spatial, temporal and causal relations. The last of these will be what plays an important role in our account of auditory sensory individuals.

In the case of the table example, the table is a composite particular because it is spatially extended. And when you look at the table, your perceptual state represents it as composite because it is spatially extended. If an object takes up some of your visual field, it is perceptually represented as spatially extended. You can attend to its top and then to its bottom, to its left side and to its right side. Seeing it entails representing these proper parts of it. In short, this would amount to your perceptual state attributing properties to a composite sensory individual, where this composition amount to spatial composition.

Most cases of visual perception are like this, hence, most of our visual states attribute properties to composite sensory individuals. But not all of them. Sometimes it can happen that an object is so far away from you so that you only see it as a dot and, as a result, you just can’t visually represent its proper parts. In this case, your perceptual state attributes properties to a non-composite sensory individual.

But even when it comes to distant dots, we can get composite sensory individuals as examples of ensemble perception show (see Ariely 2001 for a summary of the ensemble perception literature). Consider first a simple display consisting of two spatially separated dots, moving entirely synchronously with each other: as one moves left, right, etc. so does the other, each accelerating and decelerating at the same rate as the other. In this case, we attribute the property of moving left, say, to not just the individual dots, but to the composite individual of the ensemble of dots. Further, if the trajectory of these dots is slightly different, the perceptual system averages the movements and attributes this average to the ensemble. The ensemble of dots is a composite sensory individual (whose proper parts are the individual dots).

Our second example is about temporal composition. We have seen that perceiving spatially extended particulars, as long as they are perceptually represented as spatially extended (that is, if they take up more than just a dot on your visual field), counts as attributing properties to composite sensory individuals. We can make a similar claim about temporally extended particulars: perceiving temporally extended particulars, as long as they are perceptually represented as temporally extended (that is, if they take more than a split second), counts as attributing properties to composite sensory individuals.

Given that auditory stimulus is temporally extended, the sensory individual of most auditory states is composite. Again, there are exceptions that would be the analog of the dot case in the discussion of spatial composition above. If you hear a brief beep, while the stimulus does have a temporal extension, your auditory state might not represent this sensory individual as having temporal extension. But when you hear the word ‘custard’, this auditory state entails hearing the first syllable and the second one. And you can attend to the first syllable or the second. The sensory individual of most auditory states is temporally composite.

Finally, the relation between the parts of the sensory individual can also be a causal one. Here is a non-perceptual warm-up case. According to John Searle (1983), an action consists of the intention-in-action triggering the bodily movement. Action should not be identified either with the intention-in-action or with the bodily movement, but with the complex particular of the former causing the latter.[[6]](#footnote-6) So an action is a complex particular, which consist of causally related parts. As we also represent it as composite, it also counts as a composite individual.

Moving on to perceptual examples, when you see the cue ball hitting the eight ball, this is a temporally extended event and you perceptually represent it as such. You perceptually represent its parts, its beginning and its end. These parts are temporally related. But it is also an event that has two causally related parts. If we accept (for the sake of the argument) the view according to which causal relation can be perceptually represented (Michotte 1967, Scholl and Tremoulet 2000), the event of one ball hitting another is perceptually represented and it is represented as causally composite. Composite sensory individuals can be spatially, temporally or causally composite. [[7]](#footnote-7)

Our claim is that auditory sensory individuals are temporally extended, causally composite sensory individuals in exactly this sense: we attribute properties auditorily to causally composite sensory individuals. And this is a radically new account of auditory individuals inasmuch as the existing accounts of audition seem to presuppose that the sensory individual of audition is not causally composite.

While this is a new account of auditory individuals, at least one influential account of what sounds are, at least in some of its formulations, could be thought to be a special case of it. O’Callaghan considers sounds to be *medium disturbing events*: “the events in which a medium is disturbed or changed or set into motion in a wave–like way by the motions of bodies. Events such as collisions and vibrations of objects cause the sound events. Among the effects of sounds may be sound waves propagating through a medium and the auditory experiences of perceivers” (2010, p. 36; see also 2007, ch. 5; 2011, p. 395). Although the second and third sentences of this quote suggest that we should think of sounds as monadic, non-composite individuals, which are the products of vibrations and the producers of sound waves, it is, arguably, difficult to conceive of medium disturbing events without conceiving them as composites, consisting of a disturber and a disturbed. While vibrations can occur in a vacuum, and air, in the form of sound waves, can move without being immediately adjacent to an object, sounds, as O’Callaghan conceives of them, cannot occur in isolation, they require two different aspects of the environment for their existence. If this is correct, then to hear a medium disturbance is to auditorily represent a composite sensory individual.[[8]](#footnote-8)

Before moving on, it is worth emphasising that hearing a composite individual is *not* a form of mediate perception. It is not that one part of the individual is perceived *through* the hearing of another part; rather, the different parts of a composite individual are, representationally on a par. We do not think that this is a particularly controversial idea. There is no reason why a single perceptual modality cannot represent multiple types of environmental feature simultaneously. Consider, for example, Matthen’s work on sensory perception. In Matthen 2010a, for example, he suggests that an experience of a white wall looking both white, and somehow, pink in red light need not be explained by referring to two types of properties being attributed to the wall (its real white colour and its apparent pink colour, for example) but rather the wall, and *the illumination itself* being represented as separate individuals, with different colour properties being attributed to each. Or, to take an auditory example, he argues that when we hear a group of notes playing a melody, we do not hear the melody in virtue of hearing the notes (nor the notes in virtue of the melody), but represent both the notes *and*  the melody (Matthen 2010b). Where our account departs from one like Matthen’s is that separate environmental aspects are not simply represented but represented as causally related parts of a composite.

In Section 1 we outlined two debates concerning the nature of source perception: what are the sensory individuals of source hearing? How does the representation of sounds relate to the representation of sources? We also saw that, in both cases, the solutions that have been proposed face certain problems. Thinking of source hearing in terms of composite sensory individuals provides, we argue, novel answers to both these questions, which avoid the difficulties of the competing views.

## 3. Composite Individuals and the Problems

### 3. 1 Sources

Composite sensory individuals allow us to unite both candidates that have been put forward for source hearing, thereby retaining the benefits of each. We saw in Section 1 that a problem of identifying source hearing solely with the representation of vibrations is unsatisfactory because vibrations, in and of themselves, do not seem to be an aspect of the world that listeners have an interest in. However, thinking of source hearing as just the representation of interaction events themselves runs into difficulties when we consider collision events: in these cases, the objects have ceased to interact by the time we enjoy an auditory experience, something better explained by the vibration account.

We suggest that hearing a source should be thought of as the representation of a composite individual, consisting of an interaction event part and a vibration part, with a causal relation between the two. As such, an individual includes an interaction event as a part, it includes the properties of those interacting objects, and so is an item that listeners would have an interest in representing. Such an account therefore easily avoids the problems faced by the vibrations account of source hearing.

Representing a composite also allows for a novel way of dealing with collision events. We can think of the representation of interaction events and the vibrations they produce as a *temporal* composite, and the fact that the latter are represented as a part of composite which includes the former allows for the *representation* of the former. Another way of putting this would be to say that the representation of the vibrations part of the individual allows for the *post-dictive* representation of its earlier interaction event part.[[9]](#footnote-9)

### 3. 3 Sources and Sounds

Finally, the idea of composite sensory individuals suggests a novel way of linking sound hearing with source hearing. We saw in 1.2 that whereas the view put forward by Casati et al. collapsed the distinction between sound and source, on views which identify sounds as sound waves, or medium disturbances, there is a distinction between representing a sound and representing a source event. In 1.3 we saw that according to Nudds source hearing is possible because sounds are heard as bearing the property of having been produced by a source of a certain type, but that it was not obvious why representing such a property would allow for bona fide *representation* of the source; and that according to O’Callaghan we hear sounds as parts of sources, but that this approach is weakened if we consider some disanalogies between hearing sounds and seeing the front facing surfaces of objects.

We suggest that asking how we hear sources *through* hearing sounds posits a false dilemma: we do not hear the former *via* the hearing of the latter, rather, hearing consists in the representation of an individual with component parts, none of which is perceived *prior* to any other. Consider what it is to hear sources as composite individuals: hearing a source is to represent a particular consisting of vibrations and an interaction event. Neither individual is heard through the other: although the interaction event is heard as the cause of the vibrations, there is no order of perceptual priority –the parts are simply represented as causally related parts of a composite. We can think of the relation between hearing sources and hearing sounds in a similar way: hearing a sound and its source consists in representing a composite with causally related parts. We can think of the relation between hearing sources and hearing sounds in a similar way: hearing a sound and its source consists in representing a composite with causally related parts. Although this proposal involves both perceived causality and parthood, the lack of perceptual priority between its causal components allows us to avoid the difficulties that Nudds and O’Callaghan’s proposals face.

If hearing a source is a matter of representing a composite with a vibration part and an interaction part, and the relation between hearing sounds and sources is a matter of hearing the two as composite parts of the same individual, there is an obvious final step that could be taken. Hearing involves the representation of a *single* composite sensory individual, consisting of a sound part, a vibration part, and an interaction part, all represented as causally related to one another: interactions events are heard as causing vibrations, which are heard as causing sounds.

## 4. Conclusion

The resulting account has a fairly unorthodox view of what could count as auditory individuals. Matt Soteriou in a recent paper argued that we should endorse a more pluralistic approach to what we call auditory individuals (and what he calls “the bearer of acoustic properties” 2018, p. 36). Our move could be considered to be the exact opposite of Soteriou’s, albeit the picture we end up with is somewhat similar to Soteriou’s.

For us, there is only one auditory individual: a mereologically complex composite sensory individual. In different auditory episodes, and depending on our momentary interests, different parts of this mereologically complex composite sensory individual may be more salient – you can attend to various parts of this composite sensory individual – to the sound, to the vibration event, and so on. But there is no plurality when it comes to the sensory individual itself. But our approach, like Soteriou’s, opens the way to accommodating seemingly conflicting intuitions about the nature of auditory individuals – not by allowing a variety of auditory individuals, but by allowing only one, albeit a composite one.

Finally, a word about the scope of our strategy. We have been focusing on auditory individuals. But similar questions can and should be asked about the other sense modalities as well. Olfaction is an obvious target, for example, where structurally similar questions are asked about the nature of olfactory individuals: are they odors (the equivalent of sounds) or source events (see Batty 2011 for an overview)? Do we smell the odor of burnt toast or of the source event of the burning of toast? Again, one novel way of answering this question would be to claim that the olfactory individual is a composite sensory individual (comprising both of the source event and of the odor).

### References:

Ariely, Dan. 2001a. “Seeing Sets: Representation by Statistical Properties.” *Psychological Science* 12 (2): 157–62.

Aristotle (1984). De Anima. In J. Barnes (ed.), *The Complete Works of Aristotle: The Revised Oxford Translation*. Oxford: Oxford University Press.

Batty, C. (2011) Smelling lessons. Philosophical Studies, 153, 161–74.

Bregman, A. S. (1994) *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Byrne, A., and D. Hilbert. (2003) Colors and reflectances. *Behavioral and Brain Sciences* 26(1): 3–21

Casati, R., Di Bona, E., & Dokic, J. (2013). The Ockhamization of the event sources of sound. *Analysis*, 73(3), 462-466.

Choi, H. and Scholl, B. J. 2006: Perceiving causality after the fact: postdiction in the temporal dynamics of causal perception. *Perception*, 35, 385–99.

Clark,​ ​A.​ ​(2000).​ ​*A*​​*theory*​​*of*​​*sentience*.​ ​Oxford: Oxford​ ​University​ ​Press.

Clark, A. (2004). Feature–placing and proto–objects. *Philosophical Psychology*, 17(4), 443– 469.

Clarke, Thompson (1965). Seeing surfaces and physical objects. In: Philosophy in America, ed. M. Black. Ithaca, NY: Cornell University Press, pp. 98-114.

Cohen, J. (2004). Objects, places, and perception. *Philosophical Psychology*, 17(4), 471–495.

Di Bona, E. (2017). Towards a rich view of auditory experience. *Philosophical Studies*, 174(11), 2629-2643.

Dretske, Fred 1981 *Knowledge and the Flow of Information,* Cambridge, MA.: The MIT Press.

Kulvicki, J. (2008) The nature of noise. *Philosophers’ Imprint*, 8(11), 1–16.

Kulvicki, J. (2014) Sound stimulants: defending the stable disposition view. In Stokes, D. Matthan, M, Biggs, S., *Perception and Its Modalities*. Oxford: Oxford University Press, 205–221.

Kulvicki,​ ​J.​ ​(2016)​ ​Auditory​ ​Perspectives,​ ​In​ ​Nanay,​ ​B​ ​(ed.)​ ​*Current*​​*Controversies*​​*in*​​*Philosophy*​​*of Perception*.​ ​New​ ​York:​ ​Routledge.

Leddington, J. (2013). What we hear. In Brown, R. (ed.), *Consciousness Inside and Out*. Dordrecht: Springer.

Matthen,​ ​M.​ ​P.​ ​(2005).​ ​*Seeing,*​​*doing,*​​*and*​​*knowing:*​​*A*​​*philosophical*​​*theory*​​*of*​​*sense*​​*perception*. Oxford: Oxford University Press.

Matthen, M. (2010a) How things look (and what things look that way). In Bence, N. (ed.), Perceiving the World. New York: Oxford University Press, pp. 226–53.

Matthen, M. (2010b) On the diversity of auditory objects. Review of Philosophy and Psychology, 1, 63–89.

Michotte, A. 1963 *The Perception of Causality*. New York: Basic Books.

Nanay, B. (2010). Perception and imagination: Amodal perception as mental imagery. *Philosophical Studies* 150: 239-254.

Nanay, B. (2011). Do we see apples as edible? *Pacific Philosophical Quarterly* 92: 305-322.

Nanay, B. (2013). *Between perception and action*. Oxford: Oxford University Press.

Nanay, B. (2018). The importance of amodal completion in everyday perception. *i-Perception* 9 (4): 1-16. doi: 10.1177/204166951878887

Nudds, M. (2009) Sounds and Space. In M. Nudds and C. O’Callaghan (eds), *Sounds and perception: New philosophical essays*. Oxford: Oxford University Press, 69–96.

Nudds, M. (2010a) What sounds are. Oxford Studies in Metaphysics 5: 279–302

Nudds, M. (2010b) What are auditory objects? *Review of Philosophy and Psychology*. 1(1), 105–122.

Nudds,​ ​M.​ ​(2014).​ ​Auditory​ ​Appearances.​ ​*Ratio*,​ ​27(4),​ ​462–482.

O’Callaghan, C. (2008). Object perception: Vision and audition. Philosophy Compass, 3(4), 803-829.

O’Callaghan, C. (2010) Sounds and events. In M. Nudds and C. O’Callaghan (eds), *Sounds and perception: New philosophical essays*. Oxford: Oxford University Press, 26–49.

O’Callaghan,​ ​C.​ ​(2011):​ ​Hearing​ ​properties,​ ​effects,​ ​or​ ​parts? ​*Proceedings*​​*of*​​*the*​​*Aristotelian Society*​​111:​ ​375​ ​–​ ​405

O'Shaughnessy, B. (2002). *Consciousness and the World*. Oxford: Clarendon Press.

Porter J., Craven B., Khan R., Chang S-J., Kang I., Judkewitz, B., and Sobel, N. (2007) Mechanisms of scent tracking in humans. *Nature Neuroscience*, 10, 27–9.

Scholl, B. J. and Tremoulet, P. D. (2000) Perceptual causality and animacy. Trends in Cognitive Sciences, 4, 299–309.

Searle, J. R., & Willis, S. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Siegel, S. (2006). Which properties are represented in perception. in n Gendler, T. S., Hawthorne, J. (eds.), *Perceptual Experience*. Oxford University Press. 481–503. (2006).

Sorensen, R. (2008) *Seeing Dark Things*. Oxford: Oxford University Press.

Soteriou, M. (2018) Sound and Illusion. in Crowther, T., & Mac Cumhaill (eds.), C. *Perceptual Ephemera*. Oxford: Oxford University Press, in print.

Young, N. (2017). Hearing Spaces. Australasian Journal of Philosophy, 95(2), 242-255

Young, N. (2018) Hearing Objects and Events. *Philosophical Studies*, 175(11), 2931–2950

1. Leddington, a proponent of the property view, distinguishes two ways of understanding the idea that sounds ‘have’ pitch, timbre, and loudness. If sounds are particulars then ‘have’ can be understood in terms of *property attribution.* If sounds are properties, ‘have’ can be understood in terms of *constitution*; sounds have pitch, timbre, and loudness in a similar way to how colours have a certain brightness, hue, and saturation. [↑](#footnote-ref-1)
2. In particular, there is disagreement as to whether hearing a source event involves hearing the objects involved (Kulvicki XXX; Young 2018) or the events but not objects (O’Callaghan REF, Leddington REF) [↑](#footnote-ref-2)
3. The idea that vibrations are a type of source event can also be found in O’Callaghan (2008), who says: “Sounds and streams furnish information about the events and happenings – the collisions, vibrations, and interactions – that commonly make or produce sounds” (p. 824). Green also considers this way of thinking about sources which he calls the “minimal object view” (2018, p. XX). A somewhat similar view has been put forward by Kulvicki (2008, 2014) on which sounds are stable dispositional properties of objects. Specifically, the disposition to vibrate at certain frequencies. On this view vibrations are not identical to sounds, but rather reveal sound properties, in the same way that on reflectance theorists (e.g. Byrne and Hilbert 2003) would say that light reflecting from an object reveals that object’s colour property. Although, on this view, like that of Casati et al. audition involves the representation of only one sensory individual (rather than one for sound hearing and one for source hearing), Kulvicki does not identify hearing a source event with hearing an object vibrate. Rather, he suggests that “that we hear events by hearing stable properties of objects”. See Kulvicki 2008, pp. 11 – 13 for details of this view, and Young 2018 for criticism. [↑](#footnote-ref-3)
4. Given that it takes time for sound waves to propagate, and time for our auditory systems to produce a representation once they have reached our ears, the vibrations are, of course, not *completely* in sync with our auditory experience. The point here is that they are closer to being synchronous than the collision event. [↑](#footnote-ref-4)
5. Nudds (personal communication) has since moved away from this way of reconciling sound hearing with source hearing. [↑](#footnote-ref-5)
6. Needless to say, our argument about auditory individuals does not presuppose the truth of Searle’s claims about the philosophy of action, it is merely a demonstration that positing causally composite individuals is not a completely unprecedented move. [↑](#footnote-ref-6)
7. On Green’s theory of perceptual objects, our perceptual systems “function to pick out entities whose parts are causally related”, although he stops short of suggesting that such relations are represented in perceptual content. [↑](#footnote-ref-7)
8. Some remarks by O’Callaghan suggest that he is sceptical that hearing a medium disturbance must involve perceiving the medium and the disturber. He says: “one might maintain that to perceive an interaction—for that interaction to count among the things one perceptually experiences—requires perceiving its relata...While this might be the norm in the case of visible objects, nothing prevents interaction itself from having perceptual significance in absence of perceptual awareness of what interacts...one might be aware of a disturbance event through experiencing its distinctive audible qualities, while immediately experiencing neither the object nor the medium” (2007; p. 68- 69). However, one worry we might have with this is that audible qualities *supervene* on the relata. Pitch, timbre, and loudness depend on, respectively, frequency, spectral composition, and amplitude –these are properties instantiated by the vibrating object and the disturbed air, but not the causal interaction between the two. [↑](#footnote-ref-8)
9. Choi and Scholl (2006) present empirical work which they suggest shows that collision events can be perceived post-dictively. [↑](#footnote-ref-9)