Abstract of thesis entitled

"Hearing Waves: A Philosophy of Sound and Auditory Perception"

Submitted by

Kwok Ka Wing

for the degree of Doctor of Philosophy
at The University of Hong Kong
in September 2020

This dissertation aims to revive wave theory in the philosophy of sound. Wave theory identifies sounds with compression waves. Despite its wide acceptance in the scientific community as the default position, many philosophers have rejected wave theory and opted for different versions of distal theory instead. According to this current majority view, a sound has its stationary location at its source. I argue against this and other alternative philosophical theories of sound and develop wave theory into a more defensible form.

Philosophers of sound tend to emphasise how sounds are experienced to be in their arguments. Most often, it is assumed that that which appears to be a distally located bearer of auditory properties in an auditory experience is a sound. Chapter 1 argues that if this distal entity is the sound source instead, many of the existing theories of sound will be severely affected.

Chapter 2 discusses auditory perception and criticises the common assumption that we hear non-sound entities in virtue of hearing sounds. I show that this assumption begs the question against certain theories of sound and that the contrary view that sound sources can be directly heard is more plausible.

If sound sources can be directly heard, then features commonly attributed to sounds based on auditory experiences might rather be features of sound sources. I examine eight of such features in Chapter 3. Only four of them survive.

Chapters 4 and 5 review the existing theories of sound. After a taxonomy of existing theories of sound, each theory is criticised one-by-one. Some of them are problematic precisely because they rely on the implausible assumption that that which appears to be distally located in an auditory experience is a sound rather than a sound source.

Lastly, Chapter 6 focuses on wave theory. It begins with two positive arguments for wave theory in general, followed by my replies to two common objections in the literature. I then move on to develop my version of wave theory. There are two core aspects of my view. The first one is a metaphysics of compression waves; the second is an account of what it is to hear compression waves. After comparing my view with a similar theory, I demonstrate the explanatory power of my view in two steps. First, the eight commonly accepted features of sounds examined in Chapter 3 are revisited. It turns out that my view can accommodate all of them. Second, explanations for four special sound-related phenomena are offered at the end of the chapter.

I conclude in the last chapter with the suggestion that, as a philosopher, the best way to defend wave theory is to offer a better understanding of auditory perception which explains how compression waves are experienced.

(Word count: 452)

Hearing Waves: A Philosophy of Sound and Auditory Perception

by

Kwok Ka Wing

郭嘉榮

B.A., MPhil (Lingnan)

A thesis submitted in partial fulfilment of the requirements for
the Degree of Doctor of Philosophy
at The University of Hong Kong.

September 2020

To Hongkongers

Declaration

I declare that this thesis represents my own work, except where due acknowledgement is made, and that it has not been previously included in a thesis, dissertation or report submitted to this University or to any other institution for a degree, diploma or other qualifications.

Signed ……………………………………………………………

Kwok Ka Wing

Acknowledgements

Some years ago, my MPhil supervisor, Prof Paisley Livingston, lent me his copy of Casey O'Callaghan's book *Sounds: A philosophical theory* (2007). I can still remember reading it on the train and the feeling of being amused by the challenges against the received scientific view that sounds are compression waves. For some reason, I did not immediately pursue a PhD. But when I finally decided to continue my research journey, I knew that I will enjoy every minute in the next three years if I work on the philosophy of sounds and auditory perception. And time has proven me right. I should thank Paisley for introducing this fascinating topic to me.

I would also like to thank Prof Max Deutsch, my supervisor at HKU, for his encouragement and support. Although my topic is only marginally related to his expertise, his perspective as an insightful reader with a general background has been essential to my being able to take a critical stance toward the common assumptions in the field. Without him, this work might just be a dull reiteration of what has already been said in the literature, not to mention that my position would be completely different.

The final year of my study happened to be a very difficult time for people in Hong Kong. The anti-extradition movement and the COVID-19 pandemic provided perhaps the least favourable environment for abstract theoretical research like the one presented in this work. There were too many exhorted days of reading and writing papers followed by sleepless nights of desperately catching up with news and stories about what happened on the street. The end of all these is still not in view, but I know that the arguments and ideas in this work will always be accompanied by my memory of the unforgettable and unforgivable moments in history. My life in Hong Kong will never be the same. I will forever be indebted to all Hongkongers who have sacrificed more than they should. This work is dedicated to them.

My gratitude also goes to members of my examination committee: Prof Matthew Nudds, Prof Giorgio Biancorosso, Dr Jamin Asay, and Dr Jennifer Nado. Their kind and supportive comments had magically turned the supposedly most stressful two hours in these three years into the most enjoyable afternoon.

I thank my family for allowing me to do what I love to do, even when they have no idea what exactly I am doing. Lastly, I thank my peers in the Philosophy Department at HKU for the serious discussions and delightful chit-chat, as well as keeping the morale up: Wan-Shen Chan, Carrie Choi, Yun-Chak Chong, Nikolaj Bichel, Sanja Särman, and Felix Yeung (listed alphabetically by surname). Special thanks to Monica Ding for her emotional support and coffee.

Contents

Illustrations

Figure

Tables

1 Introduction

1.1 Sounds, Auditory Experiences, and the Veridicality Requirement

What are sounds? Science teaches us that they are compression waves. This wave theory, however, is challenged by some philosophers. For example, Robert Pasnau (1999, p. 311) claims that this is inconsistent with the apparent locations of sounds. Sounds seem to have stationary locations at a distance, but compression waves are not stationary. Rather, they propagate in the medium. Pasnau thinks that it is better to reject wave theory than to admit that we are constantly deceived.

Underlying Pasnau's argument is the assumption that veridical auditory experiences tell us what and how sounds are. Auditory experiences are mental states which represent the external world in a certain way. Moreover, they are more like beliefs than desires in having a mind-to-world direction of fit. An auditory experience is veridical when its content fits how the external world is. Otherwise, the experience is defective—it is illusory or hallucinatory. It follows that if a veridical auditory experience represents a sound as located at a distance, the sound is located at a distance.

Knowing what it is for an auditory experience to be veridical does not provide us with any sufficient reason to believe in our actual auditory experiences. If our experiences are never veridical, then we should not trust them. Since the nature of sounds is in question, we cannot determine the veridicality of our auditory experiences directly by checking their contents with what sounds are. A less direct method is required.

Here it is helpful to appeal to the fact that hearing is generally a reliable source of information about the external world. It alerts us to threats and opportunities in the surroundings. It would be mysterious if auditory experiences are never veridical. Indeed, the general success of our hearing-guided navigation in the world is best explained by the supposition that auditory experiences are at least largely veridical, and hence it is reasonable to accept this supposition.

This general conclusion about the veridicality of auditory perception takes us a step further in the investigation of the nature of sounds. If a theory of sound implies that our auditory experiences massively and systematically misrepresent sounds, then it conflicts with the supposition that auditory experiences are largely veridical. Therefore, we shall reject such a theory. This idea can be stated as the following constraint on theories of sound:

> *Veridicality Requirement*
> A theory of sound should not imply, or at least should minimise the extent to which auditory experiences contain massive and systematic errors.

Many philosophers of sound accept this requirement, e.g. Casati and Dokic (2010); Nudds (2009, pp. 77-78); O'Callaghan (2007b, p. 14); and Pasnau (1999, p. 311). Interestingly, they draw different conclusions about where sounds are. Casati, Dokic, O'Callaghan, and Pasnau hold that sounds are located where they appear to be, i.e. at their sources. In contrast, Nudds locates sounds in the medium. Given that they all agree that theories of sound should respect how sounds appear to be, why could their views be so different? Because they understand the phenomenology of auditory experience differently. Nudds holds that while sound sources are heard to be located at where they are, there is no sense in which sounds are also heard at the same locations. Therefore, sounds are not heard to be anywhere. In contrast, the other four philosophers above simply take the entities which appear to be located at the sound sources as sounds and hence hold that sounds are heard to be located there.

The difference between these two camps brings us to an important issue: the phenomenology of auditory experience is open for interpretation. It is unlikely that the dispute between those philosophers is caused by different auditory phenomenology. Particularly, given that they all agree that something is heard to be located at a distance, the issue concerns what that thing is. Our auditory experiences represent certain things as being in certain ways, but the identities of those things are not parts of the phenomenology. Saying that they are sounds or that they are sound sources are both substantial claims about auditory experiences which need to be backed up by arguments. Before we can evaluate theories of sound based on how sounds appear, we should therefore first carefully examine auditory perception.

Now, we can see that the application of the veridicality requirement in the philosophy of sound is called into question. To determine whether a theory of sound implies an error theory of auditory perception, we need to first determine what exactly are represented in veridical auditory experiences. If, for instance, we only hear sound sources but not sounds as having stationary locations at a distance, then, *pace* Pasnau, wave theory is not inconsistent with the apparent locations of sounds. Moreover, it is possible that sounds are not represented spatially at all, in which case the spatial content of our auditory experiences is entirely neutral regarding the nature of sounds. Instead of refuting wave theory, Pasnau's argument may just show that he and the similarly minded people have mischaracterised the phenomenology of auditory experience.

1.2 Overview

This work is an attempt to defend wave theory by criticising and correcting the flawed conception of auditory perception underlying objections against wave theory.

Chapter 2 begins our inquiry by first examining auditory perception. The main target is the widely accepted idea which I call "sonicism": we hear non-sound entities in virtue of hearing sounds. I argue that even if this idea is true, in our context where the nature of sounds is in question, we have no non-question-begging reason to accept it as a theoretical background for the further discussion of what sounds are. Moreover, toward the end of the chapter, I show that we can arrive at a simpler and more plausible theory of auditory perception if we reject sonicism. More precisely, the idea is that not only do we have no reason to think that sound sources are not direct objects of auditory experiences, admitting them to be direct auditory objects can indeed result in a better theory of auditory perception.

A consequence of this simpler view of auditory perception is that our ordinary conception of sound might have misattributed features of sound sources to sounds. Chapter 3 then continues the discussion in the previous chapter by re-examining the commonly accepted features of sounds. It turns out that four out of the eight discussed features might not really be features of sounds, and hence we should not employ them in arguing about the nature of sounds.

Having thus cleared the ground for our further inquiry into the main metaphysical question concerning what sounds are, I move on to review the existing theories of sound in Chapters 4 and 5. The current theoretical landscape is first introduced with the help of a taxonomy of theories of sound. Each theory is then criticised one-by-one in the remaining parts of the two chapters.

Lastly, Chapter 6 is where I provide my positive account of sound. In the first half of the chapter, I start with two arguments for wave theory. These arguments are intended to be adoptable by wave theorists in general. In other words, they do not rely on more than the very basic claim that sounds are compression waves and some general considerations. I then move on to respond to two major objections against wave theory. Both objections are based on a problematic view of auditory perception, or so I shall argue.

The second half of Chapter 6 turns to my wave theory. A fully developed wave theory should contain much more than the bare-bones identity statement that sounds are compression waves. It should also say more specifically what compression waves are. As a supplement, we also need to explain what it is to hear compression waves. After presenting these two core aspects of my view, I move on to show that it is superior to a similar theory which also ties sounds to compression waves but in a different way. The chapter ends with a demonstration of the explanatory power of my view. The eight alleged features of sounds are revisited and shown to be features which can be accommodated in my view. I also show how four special phenomena related to sounds can be explained.

I conclude in the last chapter with the suggestion that, as a philosopher, the best way to defend wave theory is to offer a better understanding of auditory perception which explains how compression waves are experienced.

2 Auditory Perception

Auditory perception provides us with knowledge about the external world. We can learn about objects and events in our environment by hearing them. However, compared with vision, hearing seems to be an inferior source of perceptual knowledge. One way to cash out this inferiority is to say that objects (or events) are more directly seen than heard. We are more removed from the same entities through hearing than vision. This further distance between us as perceivers and the external world in the case of hearing is sometimes described in such a way that suggests the existence of, as Kulvicki (2017, p. 83) call it, a "veil of sound". For instance, Berkeley says in the first dialogue between Hylas and Philonous that:

> … when I hear a coach drive along the streets, immediately I perceive only the sound, but from experience I have had that such a sound is connected with a coach, I am said to hear the coach. It is nevertheless evident, that in truth and strictness, nothing can be *heard* but *sound*: and the coach is not then properly perceived by sense, but suggested from experience. (Berkeley, 1713/1979, p. 39)

In the contemporary philosophy of sound, this idea of a veil of sound is often expressed as a thesis which I follow Leddington (2014, p. 327) in calling it "sonicism":

> *Sonicism*
> We hear non-sound entities in virtue of hearing sounds.

There is a wide acceptance of sonicism among philosophers (For examples: Armstrong, 1961, p. 20; Kulvicki, 2008a, p. 12; Nudds, 2001, p. 220; 2010b, pp. 295-296; 2014b, p. 344; 2015, p. 274; 2017, p. 99; O'Callaghan, 2007b, p. 13; 2008, p. 318; 2009b, p. 609; 2010b, p. 129; 2011b, p. 149; O'Callaghan & Nudds, 2009, p. 4; Sorensen, 2009, p. 126; Tye, 2009, p. 209, n. 3). Very often, the idea is presented as intuitively plausible and hence is introduced without further discussion.

Such a wide and uncritical acceptance is, however, a puzzling aspect of the philosophy of sound. Upon reflection, sonicism is too dubitable to warrant its status as a basic assumption, or so I shall argue in this chapter.

I begin in §2.1 with a brief discussion on intentionality, which allows me to distinguish between what I call "auditory objects" and "heard objects". Next, §2.2 evaluates two formulations of sonicism. The first one is expressed in terms of the "in virtue of" relation, while the second one is expressed in terms of the direct-indirect object distinction. §2.2.1 argues that in one interpretation of the first formulation, sonicism might be true. Nonetheless, we should not accept it for methodological reasons. As for the second formulation, §2.2.2 shows that it is based on the problematic assumption that sound sources are not direct auditory objects. This assumption is then rejected near the end of the subsection, where an explanation is given as to why we tend to think that sounds are heard whenever we hear anything. In §2.3, I conclude this chapter by pointing out the direction in which my theory of sound will develop. All in all, this chapter shows that in the philosophy of sound, sonicism should not be taken as a theoretical background on which theories of sound are developed.

2.1 Auditory Objects and Heard Objects

As a first approximation, auditory objects are objects of auditory experience. This is, however, hardly informative enough for our discussion. In this section, I examine the more general notion of objects of perceptual experiences. I admit that there might be some deep differences between hearing and other perceptual modalities. However, since our concern is independent of the peculiarities of hearing, I assume that the following discussion should apply to auditory objects.

2.1.1 Some General Background on Perception

Perception is a process through which a subject can acquire knowledge about the portion of the external world in her surroundings. In ordinary cases, this process produces a perceptual experience in the subject. A perceptual experience is a mental representation *about* something. In other words, it has the representational property of representing something. That thing is *an* intentional object of the perceptual experience. Here, the "object" in "intentional object" needs not be a physical object. It could also be an event, a state of affairs, a property, a fact, etc.

Suppose that there is an apple in front of me. When I look at it, I have a visual experience. If someone asks me "what is *the* intentional object of your current visual experience?" I would respond by saying that this question is ill-formed. Although it is probably the most straightforward to say that my experience is about that apple in front of me, my experience is also about the redness of that apple, the fact that that apple instantiates a certain shade of redness, the event in which an apple is reflecting light, the state of affairs that an apple is in front of me, etc. Therefore, perceptual experiences need not and in most cases do not have only one intentional object (Austin, 1962, p. 4; Crane, 2009, p. 475).

That there is a perceptual experience about something naturally suggests that there is something which is the object of that experience. According to this relational view of intentionality, a perceiving subject is connected to a mind-independent object in having a perceptual experience. It is unclear how this account can handle the case of hallucination, because the object of a hallucinatory experience does not exist and hence one of the *relata* of the intentionality relation is missing.

It seems there is a sense in which a veridical experience of *x* belongs to the same category with a hallucinatory experience of *x* but not with a hallucinatory experience of *y*, and we would expect that this is explained by the common intentional object of the first two experiences. If this is the intended sense of "intentional object", then the relational view fails to capture it, as it leads to a different categorisation. Accordingly, the two hallucinatory experiences would belong to the same category in virtue of being about the same object (i.e. nothing), while the veridical experience would belong to the category of experiences about *x*. Therefore, we should understand "intentional object" in another way.

Let us consider how an ordinary person would report her perceptual experience when she is not sure if it is veridical or not. Instead of saying "I see an apple", she would probably say "I *seem* to be seeing an apple". The qualification "seem" allows her to avoid committing to the existence of any apple in her surroundings. The report specifies what the reported experience is about by pointing to its phenomenal character. It is therefore natural to expect that the representational property and the phenomenal property of the reported experience are connected in some way.

How exactly these properties are connected is a big topic in the philosophy of consciousness. Without committing to some specific view, what I need in spelling out the non-relational idea of intentionality for my purpose is just the observation that a conscious perceptual experience has a phenomenal character tightly connected with its representational property, such that it is possible to translate talks about its intentional object into talks about its phenomenal character.

Consider the experience *as of* an apple in the above example. The idea is that we can translate the claim that its intentional object is an apple into the claim that it has a certain phenomenal character related to how apples look like. It is not the property of having an apple-ish outlook—this is a property of external objects. A phenomenal property should rather be a property of the experience. Let us call such a property “phenomenal apple-ish-ness”. It is the phenomenal property typically associated with seeing apples. Accordingly, an experience of an apple is a phenomenally apple-ish representation. Understood in this way, claims about intentional objects are in fact claims about the phenomenal characters of perceptual experiences.

Notice that this account allows that if the phenomenal character of an experience is rich enough, it can represent a specific object. For instance, a visual experience as of Jane could be phenomenally different from a visual experience as of Jane’s identical twin sister Mary, even if Jane and Mary look exactly the same. The phenomenal difference in the two experiences might then need to be contributed by a source other than their visual appearances. I leave this question open and simply assume this possibility in the rest of this work.

I do not try to determine which of the relational view and the non-relational view provides a better concept of intentionality. Such a debate in the more general philosophy of perception and consciousness is not the concern of this work. The brief and oversimplified discussion above is only intended to serve as a background for me to introduce how I am going to use a few terminologies.

2.1.2 Terminologies

Although the two views of intentional objects are presented as competing accounts, it seems they could also be treated as providing two equally acceptable concepts which serve different purposes in the discussion on perception. This subsection introduces how I am going to employ these two concepts in this work.

To begin with, I shall distinguish “perception” from “perceptual experience”. I will use “perception” solely to refer to the process which begins with a certain external object’s sending off sensory stimuli and ends with a subject’s having a perceptual experience caused by those sensory stimuli in a certain way. The qualification that the experience should be caused in a certain way is intended to exclude illusory experiences, which I shall ignore here for the sake of simplicity. By involving external objects in this way, perception is always veridical.

In contrast, a perceptual experience can be an element in the whole process of perception, but we can also have a perceptual experience in the case of hallucination. I will not argue with disjunctivists on whether veridical perception and hallucination share any common factor. I will simply use "perceptual experience" to refer to whatever phenomenally conscious states which possess all the phenomenal properties which can be found both in a veridical perception and an *indistinguishable* hallucination. The experience I have in a veridical perception and that in an indistinguishable hallucination can have different non-phenomenal properties, such that disjunctivists could be right in denying the existence of any common mental state in these two cases. I just call both experiences "perceptual experiences".

My usage of "intentional object" follows the non-relational view of intentionality. Accordingly, by "intentional objects", I do not mean any external entities. I simply take it as a convenient device to talk about the representational properties of perceptual experiences. To say that *x* is an intentional object is to say that there is a perceptual experience having the representational property of being a phenomenally *x*-ish representation. As for the relational view, it is helpful when I want to talk about the external object which causes a perceptual experience in the right way required for veridical perception. I use another term—"perceived object"—to refer to that entity. Related terms such as "perceived fact" and "perceived state of affairs" will also be used when I talk about entities in different ontological categories.

In the case of veridical perception, a perceptual experience is caused by what it represents in the right way. We can then say that the same entity is both an intentional object and a perceived object of the same experience, with the *proviso* that "intentional object" and "perceived object" are understood in the way introduced above, such that I should not be taken as switching to the relational view of intentionality. I also describe this entity as being perceived in the representational sense. In contrast, if an experience does not represent its external cause, that cause is a perceived object but not an intentional object. I describe it as being perceived in the causal sense. If, as in the case of hallucination, an experience represents something other than its causes, that intentional object is not perceived at all.

When I focus on a specific sensory modality such as vision or hearing, "visual object" and "auditory object" refer respectively to the intentional objects of visual and auditory experiences. As for the perceived object, I will call them "seen object" and "heard object" (*mutatis mutandis* for perceived facts, states of affairs, events, etc.).

2.1.3 Identification of Auditory Objects

Berkeley claims that we do not hear the coach. What we hear is, according to him, merely the sound of the coach. One might have the impression that there is no room for dispute. Just like it is obvious what is seen when there is an apple placed in front of me, it should be obvious what is heard when a coach passes by. If that is a sound, Berkeley is right. If that is a coach, then he is wrong. The question can be settled empirically.

This strategy does not work. Using our current terminologies, both the coach and its sound are heard objects, since they are causally connected to our auditory experience in the right way. Berkeley's claim should then be understood as being about the auditory object. The suggested strategy works only for perceived entities which have some established links with intentional objects. I know how a car is represented visually. I also know how a boat is represented visually. I can therefore tell whether the visual object is a car or a boat by simply attending to the phenomenal character of my experience.

This is, however, not the same case in Berkeley's example. Purely in terms of the phenomenal character of my auditory experience, when a coach passes by, there seems to be an individual which serves as the property bearer and the target of our attention. If we take this auditory object as a sound, then how would a coach be represented auditorily? Conversely, if we take it as a coach, then how would its sound be represented auditorily? Can we distinguish between an auditory representation of a coach and that of its sound?

No, we cannot. Berkeley's claim cannot be judged by matching the experience in question with known auditory representations of coaches and sounds. We do not have the required established link between heard objects and auditory objects. Of course, a coach passing by is auditorily distinguishable from a boat passing by. But when the distinction is cross-categorical between a coach and the sound of a coach, we have no idea what it is like to hear this distinction. There seems not to be any phenomenal difference between attending to a sound and attending to its source (Young, 2016, p. 84).

In other words, Berkeley's case is a different sort of problem. In an auditory experience, an auditory object is presented qualitatively. This auditory object does not, however, wear its identity on its sleeve. We may characterise the experience as representing a such-and-such sound, but this may just be a mischaracterisation. After all, we may also characterise the same experience in terms of the coach. Especially in the current context where the nature of sounds is in dispute, it is more appropriate to refrain from reading too much into the data and focus more directly on the phenomenal character itself. Our task as philosophers is then to find out which entity in the world is represented by this experience. I call this "the identification problem of auditory objects".

This task is related to our discussion of sonicism. Sonicism can be interpreted as a thesis about auditory objects. In that case, it presupposes that we have no trouble identifying an auditory object as a sound or a non-sound entity. The brief explication of the identification problem of auditory objects just given should show that we should call this presupposition into question.

I shall consider one possible response before moving on. It might be suggested that Berkeley's claim can be evaluated in light of some evolutionary considerations. The basic idea is that our auditory system is a product of evolution. It is a representational system shaped by the selective pressure faced by our ancestor species. To help a species survive, an auditory system should represent those things which are beneficial or harmful to individuals of the species. The suggestion is then that such things are sound sources rather than sounds. A hunter in the woods is killed by a tiger but not its roar. A stream, not its babble, quenches the thirst. Therefore, *pace* Berkeley, our auditory system represents non-sound entities in addition to, if not instead of, sounds.

This response faces a few problems. First, we just do not know what sounds are yet. It might well be the case that sounds are so tightly connected to their sources that an auditory system which represents sounds only is already good enough to serve its function.

Second, things which shape our auditory system need not be represented. Nudds (2015, p. 282) claims that the way we perceive sounds to be can only be explained in terms of our perception of sound sources. Assume that sounds are compression waves in the surrounding medium. Compression waves which arrive at our ears normally contain multiple frequency components. Our auditory system does not represent such components individually. Rather, it groups them according to certain grouping principles. For example, frequency components in the same group may be harmonically related, have the same onset time, share similar temporal profiles, and come from the same direction. These are all features which strongly indicate the common origin of frequency components of the same compression waves. Therefore, the way our auditory system represents sounds is shaped by sound sources.

Granted that this is the correct description of how our auditory system develops evolutionarily, it does not follow that sound sources are themselves being represented. All we are entitled to say is merely that our perception of sounds is idiosyncratic at the species level. Some other species which faced selective pressure imposed by factors other than sound sources might well possess auditory systems which group frequency components in completely different ways, or they might even not perform any grouping at all. It is implausible to conclude that individuals of such a species perceive sounds defectively simply because their auditory system works differently. Together with the previous idea that an auditory system which represents sounds only might already be good enough to serve its function, the thought that sound sources shape our auditory system does not force us to accept that they are represented.

Our discussion on this evolutionary objection to Berkeley's view is not meant to be conclusive. It leaves open whether a sound source such as a coach is represented in an auditory experience. This indecision put sonicism on a shaky ground, but this is not enough to knock it down. I shall move on to directly challenge it.

2.2 Sonicism

We saw earlier at the beginning of this chapter a quote by Berkeley. It suggests that, strictly speaking, only sounds are heard. We can only have indirect auditory access to non-sound entities with the help of, say, inference or learned association. This gives us an austere view of our auditory world.[1]

---

[1] In contrast to our usage, Young (2016, p. 6) reserves the term "sonicism" to this austere view.

In light of my distinction between auditory objects and heard objects, this austere view can have two versions: first, only sounds can be auditory objects; second, only sounds can be heard objects. Neither version implies the other one.

On the one hand, it is possible that the first version is true but the second one is false. If our perception of sounds is causally mediated, then the mediating entities will be counted as heard objects in the sense defined. However, this is compatible with our auditory system's representing sounds only.

On the other hand, it is also possible that the second version is true but the first one is false. Our auditory experience might veridically represent sounds while at the same time hallucinates something else, such that there are multiple categories of auditory objects. Even so, if our perception of sounds is causally immediate, then it can still be the case that only sounds are heard.

Is either version plausible? The first version faces the identification problem of auditory objects. Unless this problem is solved, we have no reason to accept that only sounds are represented in our auditory experiences.

As for the second version, it is false if sounds are some entities which cannot immediately cause auditory experiences. Therefore, for theories of sound which identify sounds with entities located at or near sound sources, they must reject the second version of this austere view. This means the second version is closely tied to some particular theories of sound. Since we have not yet discovered the nature of sounds, we cannot evaluate the second version at this stage.

The austere view of our auditory world is not widely accepted among contemporary philosophers of sound. They admit that we hear objects and events which produce sounds. This more permissive attitude toward auditory objects is reflected in the formulation of sonicism which describes how non-sound entities are heard. Therefore, our examination of sonicism should be done in the context of this more permissive view of our auditory world.

Note that once we allow that sound sources can be heard, the identification problem of auditory objects is even more acute. When an auditory object appears to be located at where a sound source is, does it make any sense to say that it is a sound rather than the sound source itself? In the current context where the nature of sounds is in question, there is hardly any non-question-begging way to distinguish sounds from sound sources (Kalderon, 2018b, pp. 115, 117). The reliance of sonicism on such a distinction thus makes its role in the philosophy of sound questionable at the outset. If what one theory identifies as a sound is a sound source in another theory, then the truth of sonicism in the former theory implies the falsity of it in the latter theory.

We formulated sonicism in terms of an "in virtue of" relation between the hearing of sounds and the hearing of non-sound entities. Alternatively, we can formulate it in terms of a direct-indirect object distinction:

> *Sonicism**
> Sounds are the direct objects of hearing, and all other audible entities are indirect objects.

It is uncommon to explicitly separate these two formulations. However, as we will see in the following discussion, the different formulations facilitate different ways of evaluating sonicism.

2.2.1 The "In Virtue Of" Relation

The "in virtue of" relation is widely employed in all kinds of discourse. However, it is not clear what exactly it is. It seems to have nothing to do with temporal order. Consider:

(T1) I see the moon in virtue of seeing the light coming from it.
(T2) The time-traveller kills herself in virtue of killing her grandfather.
(T3) I move my hands in virtue of moving my arms.

T1 and T2 involve events in opposite temporal sequences, and T3 involves simultaneous events.

Also, the spatial relation does not seem to be relevant.

(S1) I pet the cat in virtue of petting its back.
(S2) I bring my credit cards in virtue of bringing my wallet.
(S3) I melt the bust in virtue of melting the bronze.

S1 and S2 involve entities in converse part-whole relation, and S3 involves spatially overlapping entities.

To evaluate sonicism, we should begin by clarifying what it means to say that hearing sounds and hearing non-sound entities are related by an "in virtue of" relation.

2.2.1.1 Frank Jackson's Analysis

Let us start with Frank Jackson's analysis of the "in virtue of" relation:

> An *A* is *F* in virtue of a *B* being *F* if the application of '– is *F*' to an *A* is definable in terms of its application to a *B* and a relation, R, between *A*s and *B*s, but not conversely. This gives us an account for the indefinite case. We obtain an account for the definite case as follows: *This A* is *F* in virtue of *this B* being *F* if (i) an *A* is *F* in virtue of a *B* being *F* (as just defined), (ii) this *A* and this *B* are *F*, and (iii) this *A* and this *B* bear R to each other. (Jackson, 1977, p. 18)

Jackson understands the "in virtue of" relation as an asymmetric definitional relation between the *relata*. We might broaden it to take it as a kind of asymmetric explanatory relations.[2] Also, the explanation should appeal to a further relation R. Moreover, Jackson explicitly denies that the "in virtue of" relation in the sense relevant to his discussion of perception "stand for causal connexions or counterfactual conditions" (ibid., p. 16).

It is not clear in the quote above to what ontological category the *relata* of this relation belong. We may, however, make a reasonable guess based on one of Jackson's examples: Jackson lives in Australia in virtue of living in Melbourne. Here we do not find two instances of living. Rather, there is only one instance of living which can be described in two ways: living in Melbourne and living in Australia (ibid., pp. 18-19). The same instance of living at the same time constitutes the obtaining of two states of affairs. Since one view concerning the nature of fact is that "a fact is just an obtaining of state of affairs" (Mulligan & Correia, 2017, §1.1),

[2] Baldwin (1990, pp. 240-241) has the same suggestion. See also Bermúdez (2000, pp. 356-357) for a discussion on cashing out the "in virtue of" relation in terms of definability.

what we have here are two facts about the same instance of living. So, the *relata* of the "in virtue of" relation are facts.[3]

Let us then spell out what sonicism means in terms of Jackson's analysis of the "in virtue of" relation. Consider again Berkeley's example. Suppose a coach is driving along the street. According to sonicism, I hear the coach's passing-by in virtue of hearing its sound. I have a single auditory experience. This experience constitutes two facts: that I hear a coach's passing-by and that I hear the sound produced by this event. The obtaining of the first fact is asymmetrically explainable by the obtaining of the second fact.

This is a definite case of a sound source being heard by a person in virtue of a sound being heard by that person. Following Jackson's analysis, this means the application of "– is heard by a person *p*" to a sound source can be asymmetrically explained in terms of its application to a sound and a relation between sound sources and sounds. Obviously, this relation should be the relation "… being the source of …", i.e. a sound source is the source of a sound.

Back to the definite case, call the coach's passing-by *c* and the sound *s*. That *c* is heard by a person in virtue of *s* being heard by that person is then analysed as saying that (i) a sound source is heard by a person in virtue of a sound being heard by that person, (ii) *c* and *s* are heard by a person *p*, and (iii) *c* is the source of *s*.

Why is the explanatory relation between a sound source being heard by a person and a sound being heard by that person asymmetric? One suggestion would be that a sound can be heard without a sound source being heard, but a sound source cannot be heard without a sound being heard. The corresponding idea in the Australia-Melbourne example would be that Melbourne can be a city of no country, but Australia cannot be a country without any city. Another suggestion would be that a sound source could be heard by hearing another sound. The corresponding possibility in the Australia-Melbourne example would be one in which Jackson lives in Sydney instead of Melbourne.

---

[3] The same view can be found in the literature. Foster (2000, pp. 1-14) clarifies the claim of direct realism—as he understands the view—in terms of the constitution of facts, which is in turns spelt out as a relation in which one fact obtains *in virtue of* another fact or set of facts. Also, Martin (2017, p. 253) explicitly states that the "in virtue of" relation is a "relation holding among facts about what is seen".

We shall see a methodological reason to refrain from accepting sonicism by considering how these two suggestions may fail to account for the asymmetricality of the explanatory relation. Recall that our overarching project is to determine the nature of sounds. At the current stage, our aim is to examine if sonicism can serve as a theoretically neutral ground for any theory of sound. Therefore, from the methodological perspective, we should not accept sonicism if doing so would commit us to certain disputable theses about sound.

Let us begin with the first suggestion. It assumes that it is possible to hear a sound only.[4] Not every theory of sound is compatible with this possibility. If, for example, sounds are properties of sound sources, then hearing a sound only would be a case of hearing a property without hearing its bearer. Unless we accept this highly dubious claim, the first suggestion is therefore incompatible with property theories of sound, and hence it is not theoretically neutral.

The problem of the second suggestion concerns the disputable assumption that while a sound source could produce more than one sounds, a sound could only be produced by one sound source. The explanation is asymmetric because I could be hearing another sound *s** when I hear *c*, but I could not be hearing another sound source *c** when I hear *s*.

To see why this assumption is disputable, we can consider the possibility that sounds are compression waves. The sound source *c* in the above example is an event of a coach's passing-by. It is a particular which occurs only once, hence it cannot produce another compression wave at another time. During the period of its sole occurrence, the sound source produces only one compression wave which spreads through the surrounding medium.[5] Therefore, there is only one sound *s* produced by *c*. It is impossible to hear *c* by hearing another sound *s**. The second suggestion therefore cannot explain the asymmetricality of the explanatory relation.

What I have said is, of course, not decisive against sonicism when it is understood in terms of Jackson's analysis of the "in virtue of" relation. For one thing, I have not shown that there cannot be any theoretically neutral explanation of asymmetrical explanatory relation between hearing a sound and hearing a sound source. However, I cannot see what such an explanation would be. More generally, it is

[4] This possibility is further discussed in §2.2.2.2.

[5] I argue for this metaphysics of compression waves in Chapter 6, mainly in §6.5.

unclear how to evaluate sonicism without spelling out the relation "… being the source of …" in greater detail, but different theories of sound would substantiate this relation differently. It is, therefore, unclear if we can evaluate sonicism without tacitly committing to some theory of sound. I deem it the sonicists' burden of proof to provide a theoretically neutral explanation, not only in the case of sound sources, but also in every other case of audible non-sound entities. Until then, we should suspend our judgement regarding sonicism as understood here for methodological reasons. Sonicism might turn out to be true, but this is something we can ascertain only after the nature of sounds has been found.

2.2.1.2 Causal Analysis

Jackson's analysis of the "in virtue of" relation does not get us anywhere closer to a clear idea of how sonicism could be evaluated. Since Jackson stresses that his analysis is not causal, we might wonder if a causal analysis could do better.

One example of the causal analysis can be found in Casati, Di Bona, and Dokic (2013, p. 465), which is based on the distinction between direct and indirect awareness offered by Huemer (2001, pp. 55-56):

> In general, you are indirectly aware of *x* if you are aware of *x*, but your awareness of *x* is based on your awareness of something else.

Huemer adds that one feature of this "based on" relation is that it is causal. Therefore, the interpretation of the "in virtue of" relation offered by Casati et al. is also a causal one.

Consider the causal *relata* first. We saw in Jackson's analysis that the *relata* of the "in virtue of" relation are facts. We might assume that this is also the case for the causal analysis. Accordingly, using the coach example, the cause is the fact that I hear the sound of the coach's passing-by, and the effect is the fact that I hear the coach's passing-by.

We ignored the distinction between auditory objects and heard objects when we discussed Jackson's analysis of the "in virtue of" relation. This is because our criticism is independent of this distinction. However, for the causal analysis, we should bring this distinction back into our discussion.

If we interpret the two facts as facts about heard objects, then sonicism can be true only if sounds immediately cause our auditory experiences in veridical perception. Therefore, this interpretation begs the question against theories of sound which identify sounds with mediated causes of our auditory experiences. We should not accept it at this stage in our study of the nature of sounds.

If we instead interpret the two facts as facts about auditory objects, then sonicism can make sense only if we can solve the identification problem of auditory objects. In our previous discussion of this problem, we assumed that a representation of a sound and a representation of the source of that sound are two representations. One possible solution to the problem is to reject this assumption. We have one auditory experience, and this experience constitutes two facts: that it represents a sound and that it represents the source of that sound.

We mentioned in §2.1.1 that an experience can have multiple intentional objects. So, it should be fine to say that an experience can represent both a sound and its source. The problem concerns how these two representational facts could be causally related.

To begin with, notice that this should be a case of simultaneous causation. If the cause and the effect are not simultaneous, then there should be a time at which the experience represents the sound only, and at a later time, the experience becomes a representation of the sound source as well. However, if there is such a difference in the representational content, then there should be two experiences—one representing only a sound, and another one representing both a sound and its source. The first experience is irrelevant here, as we are concerned with cases where both a sound and its sources are represented. Focusing on the second experience, the cause and effect are simultaneous.

However, if the two representational facts obtain simultaneously, then why should we believe that they are causally connected? Why would they not be caused by a common cause without themselves being causally connected? A possible response is to assume the counterfactual theory of causation and show that if the sound was not represented, then the sound source would not be represented. Generalising from the case of sound sources to every other case of non-sound entities, this strategy needs to show that nothing other than sounds can be auditorily represented on its own. Since I understand the representational property of representing an *x* as the property of being a phenomenally *x*-ish representation, we can approach this question by examining the phenomenology of auditory experiences.

2.2.2 Direct Auditory Objects and Sound Sources

It would be easier to see how to evaluate sonicism by examining the phenomenology of auditory experiences if we consider the alternative formulation of sonicism in terms of the direct-indirect object distinction. Recall that:

> *Sonicism**
>
> Sounds are the direct objects of hearing, and all other audible entities are indirect objects.

We might separate sonicism* into two claims. The positive claim is that sounds are directly heard. The negative claim is that non-sound entities (e.g. sound sources) are not directly heard.

I assume the direct auditory object in an auditory experience appears as a distally located individual which bears auditory properties. For sonicists, this assumption just means that sounds appear to be distally located. This is not an indisputable assumption, as not every philosopher of sound agrees that sounds appear to be so located. Nonetheless, not much in the following discussion would be affected by this disagreement, and I believe the same conclusion can be reached even if the direct auditory object does not appear to be distally located.

Sonicism* holds that the direct auditory object is a sound rather than a non-sound entity. This can be taken as a response to the identification problem of auditory objects. To evaluate this claim, I shall focus on the more specific negative claim that sound sources are not direct auditory objects, which is equivalent to the claim that auditory experiences do not have the property of being a phenomenally sound-source-ish representation. This means we can evaluate this claim by examining if there is any reasonable phenomenological ground to say that direct auditory objects do not appear like sound sources. We shall consider three such possible grounds.

The first possible ground is that sound sources do not possess the properties direct auditory objects appear to have. We hear pitch, loudness, and timbre. But these auditory properties do not seem to be properties of ordinary objects like desks.

When I strike my desk, I can produce a sound with pitch, loudness, and timbre. Such auditory properties are attributed to the sound, rather than the desk. Since it is the sound instead of the sound source which have the auditory properties possessed by the direct auditory object, we should not identify the direct auditory object with the sound source.

The second and third possible grounds correspond to the two observations described by Nudds (2010b, pp. 295-296) about how auditory experiences can be misleading. The second one is that sometimes we hear sounds only, but we never hear sound sources without hearing their sounds. The third one is that experiences of different sound sources can have the same auditory object. We take ourselves as hearing something again when we listen to an audio recording. The idea that we re-hear something implies that that thing can be the object of both a live experience and a later experience of the recording. In the case of musical performance, the sound source of a live performance is a musical instrument, while the sound sources of the playback are loudspeakers or earphones. Since both experiences have the same auditory object, it cannot be the different sound sources experienced.

I will show that none of these grounds is plausible. I end this section by arguing for the contrary claim that sound sources are direct auditory objects. Even if we do not reject the positive claim that sounds are direct auditory objects, sonicism* and hence sonicism understood in terms of the causal analysis of the "in virtue of" relation should still be rejected.

2.2.2.1 Auditory Properties and Sound Sources

For the first possible ground, we should ask why auditory properties cannot be attributed to sound sources. Compare a visual case. We attribute visual properties like brightness to light, but this does not forbid us from attributing the same properties to light sources. Is the sun bright? Sure! If sound sources can be directly heard, would attributing auditory properties to them commit us to any absurdity?

A possible reply is that a sound source is not stimulated all the time. During its silent moments, it is not pitched. This reply makes a point. A clarification is needed then. Sound sources indeed consist of two categories of entities: objects and events. An object would not make any sound if no event happened on it. The reply we are considering here urges us to draw this distinction explicitly. I follow Casati et al. (2013, p. 462) to label these two categories of sound sources "thing sources" and "event sources" respectively. Since an object would never be counted as a thing source if it does not take part in any event source, things sources are sound sources only in a derivative sense. The reply above shows that auditory properties cannot be attributed to thing sources *simpliciter*.

How about event sources? Notice first that different events are to be counted as event sources under different theories of sound, and even proponents of the same theory might not agree on the mereology of event sources. Consider the case of tapping on a table. Candidates of the event source may include the tapping, the vibration of the top of the table, the table's disturbance of the surrounding medium, as well as some larger events which encompass some or all of the above smaller events. No matter what event sources are, one thing is certain: there is no event source during the silent moments of a thing source. Therefore, we are rightly reluctant to attribute auditory properties to any event sources in such cases for a trivial reason—there are simply no event sources available. In contrast, when there is an event source, at least for some of the candidates listed above, we can easily identify the physical correlates of pitch, loudness, and timbre in that event. More specifically, the event involves some periodic movements.[6] Such periodic movements are composed of a rich profile of frequency components with various intensities. Roughly speaking, the profile of frequency components correlates to pitch; the intensity explains loudness; timbre is a property which systematically covaries with both the profile composition and the intensity of its components.

I do not mean that such a correlation between direct auditory objects and event sources can settle the issue. After all, there are other entities involved in the auditory perceptual process which bear similar correlations with direct auditory objects. Candidates include the compression waves in the surrounding medium, as well as the vibrations of our peripheral auditory apparatus such as the eardrums. Nonetheless, the existence of a relevant correlation should at least show that event sources have the right properties to be counted as direct auditory objects.

Furthermore, we are now able to see a problem in the first possible ground. For it to support the alleged distinction between direct auditory objects and event sources, the first ground requires two presuppositions. First, there is a distinction

[6] I ignore sounds which involve non-periodic movements for the sake of simplicity.

between sounds and event sources. Second, it is more appropriate to identify direct auditory objects with sounds than with event sources.

The first presupposition may look fair enough. Sounds are caused by event sources, and causes are distinct from their effects. However, as we will see in §4.1, there is a theory of sound which identifies sounds with event sources. As for the second presupposition, we do not have the resources to assess it at this stage of our inquiry. We still do not know what sounds are. As a result, we also have no idea what difference(s) between sounds and event sources could make the former a better candidate for our direct auditory objects. All we can say at this moment is that the second presupposition would not be acceptable before we have determined what sounds are. Arguments for various theories of sound should not be based on a prior acceptance of sonicism*. That simply begs the question.

2.2.2.2 Bare Sounds

Grant it for the moment that whenever we hear a sound source, we also hear its sound. We should question whether we do ever only hear sounds sometimes. Armstrong wants to say that we do:

> But we can be said to have heard the coach only because we have heard the sound. We may not have paid much attention to the sound, we may have been much more *interested* in the coach than in the sound, but we must have heard the sound in order to hear the coach. But the reverse implication does not hold. Somebody who heard a noise, which was in fact made by a coach, but who was unfamiliar with the noise that coaches make, could not say that he heard a coach. Or at any rate he could not say that he knew he was hearing a coach. (1961, p. 20, italics in the original)

Although he pulls back a bit at the end of the quote by limiting his denial to the hearer's self-knowledge of his auditory experience, Armstrong seems to be more attracted to the stronger conclusion that the hearer does not hear the coach.

Armstrong's view is challenged by Jackson:

> But the reverse implication *does* hold. If I hear 'a noise, which was in fact made by a coach', then *ipso facto* I hear the coach—whether or not I am in a position to *say* that I do, or *know* that I do. (1977, p. 8, italics in the original)

Jackson charges Armstrong for confusing perception with belief about perception. This is related to Armstrong's way of distinguishing between immediate and mediate perception—another way to call the distinction between direct and indirect perception. For him, a mediate perception involves inference from the immediate object. Right after the above quote, Armstrong continues with an elaboration of what he means by "inference" (op. cit., p. 21). He makes it very explicit that what he means is more like the association of ideas, which is a kind of psychological process which *suggests* the mediate object.

We may then interpret Armstrong as saying that if a perceiver is unfamiliar with coaches, the sound he hears would not be able to suggest the idea of a coach to his conscious awareness. If something does not appear in the perceiver's conscious experience, then it would be a mistake to say that he perceives that thing.

Much more has to be said to make this view clear and precise, but we can already see in this primitive formulation that—without subscribing to Armstrong's view which ties perception and belief-formation together—the basic idea in the quote above can be developed into a view which escapes Jackson's charge.

Mediated perception, accordingly, can give rise to non-belief mental representations which partly constitute our phenomenal consciousness. However, such representations are acquired through prior experiences, such that the lack of prior experiences implies the lack of the relevant representation. In that case, there would not be any such representation to constitute our conscious experience, and hence the object would not be represented in the experience. If an object is not represented in our experience, then we do not perceive it.

My aim is not to defend such a simple view of perception. My purpose in presenting it is to bring out a fundamental difference between Armstrong's conception of perception and Jackson's one. While Armstrong emphasises the representational aspect of perception, Jackson emphasises the causal aspect instead. This echoes our distinction between auditory objects and heard objects.

The distinction between auditory objects and heard objects helps us to see that Armstrong and Jackson might be talking past each other. For Jackson, to perceive a material object is to be in a certain perceptual state caused by the action of that object (op. cit., p. 1). Since the coach plays such a causal role, the perceiver hears it regardless of how he thinks of the experience. In our terms, the coach is a heard object. In contrast, Armstrong's view can be understood as saying that the coach is not represented in our auditory experience if we have no relevant prior experiences. He is then denying that the coach is, in our terms, our auditory objects in such a case.

The relevance of this detour into the more general issue about the nature of perception to our current target issue—whether we sometimes hear sounds only—is that it depends. Every sound has its cause. If it is the causal aspect of perception which matters, then we never hear sounds only.[7] If we instead focus on the representational aspect of our auditory experiences, then it may seem plausible that only sounds can be heard when we were infants, as the lack of prior experiences limits the scope of what objects can be represented in auditory experiences.

It is reasonable to say that we should consider both aspects in our proper understanding of auditory perception, but the problem then is what should we say when we try to reconcile the two views? We either do sometimes hear sounds only or do not. If we allow sound sources to be counted as heard objects, then we never hear in the causal sense sounds only. But shall we make this allowance? On the other hand, even if we accept that sound sources are not represented in auditory experiences of infants, this does not tell us what significance this case has on our—adults'—auditory experiences.

[7] But how far can we get to? Do I hear your parents when I hear your voice? That sounds absurd, but why? This is a version of a further issue about perception—the "stopping problem" (Montague, 2016, p. 19)—which concerns what can be heard in the whole causal chain leading to a perceptual experience. If there is no further constraint, then it seems we should say that we also hear events like the reflection of compression waves by our torso and pinnae and the vibration of our eardrums. This would be a very extreme if not absurd view. Therefore, we would then like to introduce some limitations from considerations about what is represented in our conscious auditory experience. Perception is neither a purely causal matter nor a purely representational matter.

No decision could be made without a more thorough investigation into the general nature of perception, which we do not have space to do. At the bottom line, we should not take it for granted that we sometimes hear sounds only.

2.2.2.3 Indistinguishable Auditory Objects

The third possible ground for the distinction between direct auditory objects and event sources has a more obvious problem—it conflates qualitative indistinguishability with identity. To illustrate this problem, we can consider a visual case.

> *Surfacism**
> Surfaces are the direct objects of seeing, and all other visible entities are indirect objects.

Obviously, surfacism* is constructed as a visual counterpart to sonicism*. We shall investigate it with a concrete example. Suppose that there is an orange on my desk. According to surfacism*, I can only directly see the part of its surface facing me. Since the facing surface is distinct from the orange as a whole, the orange itself is not directly seen. I can only see the orange indirectly.[8]

Why say so? The most plausible reason would be that only the facing surface of the orange contributes to the phenomenal character of the visual experience. Any changes to parts of the orange other than its facing surface would not have any effect on what it is like to have that experience. This shows that those other parts are not seen. Therefore, only the facing surface is directly seen. Since it is not crucial to the case that an orange is used as an example, we can generalise this case to other visible objects.

A background assumption of the above reasoning is that it is what you directly see which determines the phenomenal character of your experience. If something can be changed in whatever way without leading to a slight difference in your experience, it just shows that that thing does not determine the phenomenal character of your experience, and hence it is not directly seen.[9]

---

[8] See Fish (2004) and Martin (2007, 2017) for similar discussions.

[9] Martin, for example, explicitly holds this view:

Granted this assumption, nonetheless, it still does not automatically follow that the orange is not directly seen. Consider another way to describe the orange example. We can keep on agreeing that changing the back of the orange does not change the phenomenal character of the experience, but what if we change its frontal part? The phenomenal character of the experience will change as well. Since changing the frontal part is one way of changing the orange, the orange determines the phenomenal character of the experience. This shows that the assumption that the phenomenal character of a visual experience is determined by the things directly seen is consistent with the orange's being directly seen.

To conclude that only the facing surface of the orange is directly seen, we need to argue from the unchanged phenomenal character to the conclusion that that which is directly seen is also unchanged. We might suggest another principle that the phenomenal character of our visual experience determines what is directly seen. However, this is obviously false.

There are many things in the world which are numerically distinct but qualitatively indistinguishable. Think about the mass-produced commodities you can buy in stores. They would induce phenomenally indistinguishable experiences, but this just means that the phenomenal character cannot determine which particular object is directly seen.

We might try to save the suggested principle by denying that we directly see particular physical objects. An extremely implausible option is to say that we can only directly see abstract entities such as object-types. A perhaps less questionable option is to accept some version of sense-datum theory and say that we see numerically different but qualitatively indistinguishable objects in virtue of seeing numerically identical sense-data.

> Our primitive idea of what it is for something to be seen, and for it to look a certain way, is for it to fix the way one then experiences, that is, the phenomenal nature of one's experience. (2007, p. 707)

> … the direct objects have a constitutive role of fixing, or anyway (for intentionalists) mark the constitutive role of that which fixes, the way one's current visual experience is. (2017, p. 259)

How exactly such a theory can be spelt out needs not concern us here. The orange example is supposed to be a neutral example which motivates surfacism*. If the only way to understand this example as favouring surfacism* assumes a questionable principle which commits us to some specific theory of perception, then the example fails its job. We should better look for other ways to rescue surfacism*.

Here is our second attempt. We have seen above how it can be argued that the orange can be directly seen in a way parallel to how it was initially argued that only the surface is directly seen. Both cases involve changes to the supposed direct objects of seeing. We might then wonder if proponents of surfacism* can give a better argument based on a case which does not involve changes to the objects.

Let us then consider two objects: Object o_1 is just the same orange in the original example, and Object o_2 is a particle-to-particle replica of the frontal part of o_1, such that both objects are indistinguishable if we look at them in the front. Suppose I am now sitting at my desk and looking at o_1. I am then asked to close my eyes for 10 seconds. During that period, my colleague carefully replaces o_1 with o_2. When I open my eyes, I think I am still seeing o_1, as both objects look the same from my perspective, but I am seeing o_2 instead.

Since both experiences have the same phenomenal character, and we have no reason to suppose that either experience involves a perceptual illusion,[10] we should, therefore, conclude that the same object is seen both when I see o_1 and o_2. The only thing shared by o_1 and o_2 is their facing surface, so I see that surface in both experiences. Moreover, since the phenomenal character of my visual experience is determined by what is directly seen, so we can conclude that in both cases I directly see the facing surface.

Shall we accept this argument? One problem is that it is misleading to say that the two experiences—one of o_1 and one of o_2—have the same phenomenal character. We are only entitled to say that they have indistinguishable phenomenal characters, but there are at least two ways in which perceptual experiences are indistinguishable phenomenologically—we might be seeing the same object, or instead seeing distinct but indistinguishable objects.

The argument above assumes that the former is the case in the crucial step in which it moves from the indistinguishable phenomenal character to the sameness

[10] The mistaken belief which takes o_2 as o_1 is not a perceptual error.

in the object directly seen. However, the example clearly specifies that o_1 and o_2 are distinct objects, although they are indistinguishable from a specific perspective. Therefore, we should only say that we have two subjectively indistinguishable experiences of two perspectivally indistinguishable objects. This is perfectly consistent with the claim that both o_1 and o_2 are directly seen.

We have so far not yet identified any plausible argument for surfacism* in terms of the direct-indirect object distinction. The considerations we examined can equally well, if not better, support the view that ordinary physical objects are directly seen. This conclusion can be extended to the case of sonicism*.

Recall Nudds's observation that auditory experiences can be misleading regarding the sources of the sounds we hear. The idea seems to be that the same auditory experience can be produced by different sound sources without being illusory about the sounds. The case is then parallel to the example where we have an orange and a particle-to-particle replica of its frontal part.

Suppose I listen to my friend's live performance of a violin sonata in a concert hall, and later listen to a recording of her performance. Suppose also that the recording sounds exactly the same as how the performance sounds to me in the concert hall.[11] Therefore, my auditory experience of the live performance and that of the recording are qualitatively the same.

On neither occasion do I suffer from any perceptual illusion. When I listen to the playback with eyes closed, I might mistakenly *think* that I am hearing the live performance again. This is, however, a post-perceptual error.

As we have granted that the phenomenal character of my auditory experience is determined by its direct object, this suggests that I directly hear the same thing on both occasions. Since the live performance and the playback of the recording are two different events, they could not be that which I hear directly. In contrast, the sound of the live performance and that of the playback are the same. So, I directly hear the sound, and the performance and the playback are heard indirectly.

Again, this line of thought can be criticised for conflating indistinguishability with identity. The two experiences are subjectively indistinguishable. The two events—the live performance and the playback—are different, but they can

[11] Similar to vision, hearing is perspectival. Therefore, we need to assume that the perspectival properties of the performance are reproduced perfectly in the playback event.

nonetheless be auditorily indistinguishable. The auditory indistinguishability of the two events suffices to account for the indistinguishability of the two experiences. Therefore, the example does not provide us with any sufficient reason to think that the two events are not directly heard.

This negative conclusion regarding Nudds' observation is far from satisfactory. First, it does not positively show that we do directly hear non-sound entities. Second, even if we directly hear non-sound entities, we should still ask whether we also directly hear sounds. Third, if we directly hear both non-sound entities and sounds, then it seems our auditory experiences have a kind of systematic error. When I hear, say, a violin solo, my auditory experience does not appear to be about two distinct entities—a sound and its source. Rather, the phenomenal character of my auditory experience is that there appears to be only one individual.[12] All the auditory properties are attributed to this object. If I directly hear both the performance and the sound, then these two entities are integrated into a single auditory object in my experience. This leads to an error theory of auditory perception that our auditory experiences systematically misrepresent two entities (a sound and its source) as one individual.

Besides, it is not clear how the auditory properties of that single auditory object are determined. Are they jointly determined by the performance and the sound? Or are they determined separately, such that some are determined by the performance and the others are determined by the sound? Or are they overdetermined? More generally, it remains unexplained how we could hear two entities—a sound and its source—when there appears to be only one in an auditory experience. I shall call this problem "the unification problem".

The unification problem is a general problem for every theory of sound. More generally, it is the problem of explaining how two (or maybe more) objects of perception can appear as one individual in a perceptual experience. Indeed, we seem to have the same problem in the case of vision. We see light sources. We also see light. However, when I direct my gaze toward a light bulb, there seems to be only one thing in my visual experience. Is it the light or the light bulb?

---

[12] In reality, there might be objects which, for instance, reflect or (partially) block the compression wave. It is plausible that they might thereby enter our conscious awareness. However, here I focus on the idealised case where only the sound source and its sound—whatever it is—are experienced.

The unification problem can be solved only after we have determined the nature of sounds. For the moment, we should notice that it intensifies the identification problem of auditory objects. When I listen to an audible event,[13] there seems to be only one individual in my auditory experience. Initially, the identification problem concerns the indeterminacy in determining whether that individual is the audible event or its sound. The unification problem complicates the picture by adding a third option that that individual might be a unified representation of both the event and its sound.

Apart from these further issues, I shall consider a possible objection. It might be objected that qualitative indistinguishability is sufficient for the identity of auditory objects. Perhaps auditory objects have a different nature from ordinary objects. While it is true that different ordinary objects can be qualitatively indistinguishable, auditory objects are individuated entirely qualitatively.

Our distinction between auditory objects and heard objects is relevant here. The proposal we are discussing is that event sources are not direct auditory objects because experiences of different event sources can have the same direct auditory object. This can be understood in two ways.

First, different experiences can share the same heard event. Recall the example of live performance and its audio recording. The proposal understood in this way says that on both occasions, the causal chains involve the same entity heard by the subject. This should then be a repeatable entity. Moreover, the directness condition requires it to be located somewhere in the causal chain between the subject and the events. Some natural suggestions for this entity would be a certain property of the events, an abstract waveform of the longitudinal waves, etc. However, whatever the supposedly repeatable entity is, the proposal still cannot escape the charge of conflating indistinguishability with identity.

The second way to understand the proposal is to say that different experiences can share the same auditory object. What this means is just that the two experiences represent the same object directly. The directness condition has an important role to play. Without this condition, it would not be controversial that different experiences can represent the same object. An audio recording is a

[13] An audible event needs not be an event source—it may be a larger event which has an event source as one of its proper parts.

representation of the recorded event. Therefore, an experience of a live performance and an experience of a recording of the same performance are both representations of the same performance, but this performance is represented less directly in the second experience than in the first one.

This is, however, not what people have in mind when they claim that non-sound events are not direct auditory objects, as this allows the non-sound event to be the direct auditory object in the experience of the live performance. Therefore, this can only show that non-sound events are sometimes not direct auditory objects. In contrast, the intended claim is the much stronger one that non-sound events can never be direct auditory objects. Accordingly, the performance is indirectly represented even in the live experience. There is something other than the non-sound events which is represented directly in both experiences.

Since the intended claim should cover cases of veridical auditory perception, the direct auditory object of the two auditory experiences we are considering should then be one of the heard objects. Unless either of the two experiences are not veridical, otherwise we should conclude that the direct auditory object should be some repeatable entities. There are doubtlessly veridical experiences of live performances, so we should focus on experiences of audio recordings. Are our experiences of audio recordings necessarily illusory?

A tempting suggestion is to say that an audio recording misleads us into thinking that we are hearing a live performance. While we should allow that there can be auditory *trompe-l'œil*, this possibility can only be exceptionally rare. In most ordinary cases of audio playback, we do not only know but are auditorily aware that we are just hearing an audio recording.

For the sake of convenience, let us then consider the case of vintage vinyl records. We are in no way under any illusion that what we hear is a live performance. Therefore, we have a pair of experiences—one of a live performance and one of a vinyl record—which are both veridical. Moreover, since the experience of the vinyl record has some distinctive features, the two experiences are distinguishable. Hence, the charge of conflating indistinguishability with identity does not apply here.

Now, the claim in question is that there is some repeatable entity that is the direct auditory object in both the experience of the live performance and that of the audio recording. This claim immediately runs into problems: if the two experiences are distinguishable, where can we locate their difference?

If we locate it in the direct auditory object, then since the claim in question is that the two auditory experiences share the same direct auditory object, the repeatable entity should be able to possess different properties at different occasions. Also, despite its difference in properties, our auditory system should be able to represent its identity. When I listen to the audio recording of a live performance I attended previously, I do not only recognise the performance in the recording, but also a repeatable entity which I encountered before. My auditory system represents that entity as being the same as the one I experienced during the live performance. It is not clear how to make sense of this suggestion.

We are comparing an experience of a live performance with an experience of the playback of a vinyl record. If the vinyl record is played through a loudspeaker, then we might think that there are certain candidate entities, such as the compression waves in air, the thing sources' disturbances of the surrounding medium, etc. If such candidates are non-repeatable particulars, then we might identify the repeatable entity with, say, the abstract pattern exemplified by such periodic events. However, this is quite counterintuitive, as perception seems to be concerned with concrete particulars rather than abstract entities.

Besides, there can be playback events which do not involve any entities like those present in the case of live performance. For example, the audio recording may be played through a bone conduction headphone. No airborne compression wave is generated. The headphone transmits the vibration directly to your skull, while in the live performance the musical instrument disturbs the surrounding air. The causal chains involved in the respective experiences do not share any extracranial entity of the same kind. In this example, at most we could identify the direct auditory object with some abstract periodic pattern which is not restricted to any form of instantiation. If we can identify the repeatable entity only at this level of abstraction, the resulting conception of auditory perception is very implausible. It is much more plausible to claim instead that we have access to that abstract pattern via abstraction from what our auditory experiences represent directly.

Another option is that the differences between the two experiences lies not in the direct auditory object. Since our current example involves an experience of a vinyl record, the difference is clearly auditory, and hence we can exclude the possibility that we merely learn from the context whether we are listening to a live performance or an audio recording.

Perhaps we know the difference from what we hear indirectly. We are thus denying that the direct auditory object in both experiences are represented differently. But it is then very puzzling how there could be any *auditory* difference between indirect auditory objects when there is no *auditory* difference in the direct auditory object. This seems to be suggesting that we can bypass the direct auditory object in accessing the difference between the indirect auditory objects. If so, then it seems we have direct access to that difference. I can see no way to make sense of the idea that the difference between indirect auditory objects can be directly accessed. The direct-indirect distinction simply falls apart. Without this distinction, there is no point to maintain that event sources are not direct auditory objects.

2.2.2.4 Directness as an Epistemological Notion

We have examined and rejected three possible grounds for the distinction between direct auditory objects and sound sources. Without this distinction, the negative claim of sonicism* that sound sources are not directly heard is highly questionable. Where does the initial plausibility of this claim come from?

Consider this classic example of illusion—a stick half-immersed in water. We know that the stick is straight, but it looks bent anyway. There seems to be a divergence between reality and appearance. Since we can see the apparent bentness but not the real shape of the stick, this suggests that the appearance intercepts our perceptual access to the reality. The appearance is then considered as more directly linked to us than the stick itself does.

In this example, it is obvious that the directness of perception is playing an epistemological role. I can easily be wrong about the shape of the half-immersed stick, but I cannot mistake its appearance. If I uncritically accept what my experience presents to me, then I will form the belief that I am seeing some bent object. This belief is wrong if the thing I see is the stick. However, it seems this belief is true in some other sense. The apparent bentness is unresponsive to our knowledge about the real shape of the stick. Its incorrigibility suggests—correctly or not—that there is a sense in which my uncritical belief is true. There should really be some bent thing that we see. The thought is then that the more direct our perceptual access to an object is, the more secure our perceptual belief about that object is.

This epistemological difference between our beliefs about our experiences and beliefs about objects was arguably one of the main motivations behind sense-datum theory. The difference between those beliefs were then cashed out as a difference between the objects of the relevant experiences.

This is, however, mistaken. That which incorrigibly *looks* bent and the stick which *is* straight are taken as distinct perceptual objects, but this misconstrues the situation. The stick *is* that thing which looks bent, and this is dictated by the laws of physics. It is *the* way a straight stick looks when half-immersed in water. How should my experience represent a straight stick half-immersed in water? It should represent the stick bent-ly.

If I believe that the stick is bent, I mistakenly take the phenomenal bentness of the experience as representing physical bentness. The mistake lies not in the representation—the experience—itself but in my judgement: I conflate the phenomenal property of the experience with the physical property of the represented object. The phenomenal bentness is a property of the experience. It can be an object of introspection but not an object of perception. In contrast, the physical straightness is a property of the stick. It is an object of perception. Indeed, it is a *direct* perceptual object, as it is not mediated by other perceptual objects. Therefore, there is no perceptual object which *looks* bent apart from the stick which *is* straight.

The difference between the phenomenal properties of a representation and the represented physical properties is a subtle issue. A painting with only black and white paints on it might or might not misrepresent a coloured scene. It misrepresents only if it is a colour painting which happens to have only black and white paints on it. However, if it is a black-and-white painting, then the way it is *is* how it represents a coloured scene. The property of being covered with only black and white paints is a surface property of the paintings. Phenomenal properties are analogous to surface properties in the sense that a phenomenally bent representation might represent something as bent or something as straight. We need to consider contextual factors in deciding which case it is.

The same phenomenal property can represent different physical properties. Similarly, the same physical property can be represented by different phenomenal properties. A colour painting and a black-and-white painting can represent the same scene, but they have different surface properties. Analogously, the same object can be represented by experiences with different phenomenal properties. A person with normal colour vision and a colour-blind person can see the same object, but their experiences have different phenomenal properties.

Colour paintings and black-and-white paintings represent objects in different manners. There are still other manners of representation. A scene can be depicted pictorially. It can also be described in words. Representations with very different manners of representation normally have very different properties. However, the previous lesson still applies—they may nonetheless represent the same object. In the case of mental representations, it is also the case that the same object can be represented in different manners, such as doxastically or perceptually. The latter one can be further divided into visually, auditorily, tactually, etc.

Putting one half of a stick in water is doubtlessly not the ideal way to see its physical shape. We learn from experiences the distinction between optimal and suboptimal situations for perceptual observations. The existence of such a distinction reflects the limitations of our perceptual abilities. In different situations, the same object is experienced in different ways. All these experiences are, nonetheless, all about the same object. It is a mistake to think that different objects are represented just because our experiences have different phenomenal properties. Thinking in this way is just the same as saying that my visual experience changes its object just because I take my glasses off.

What motivates the negative claim that sound sources are not directly heard seems to be along the same line. We know that a sound source has such-and-such properties, but that thing we hear does not appear to have those properties. Instead, it is represented as pitched, loud, and having a certain timbre. Therefore, what we hear is not the sound source but a merely auditory object—a sound. I suspect this may be a possible origin of the ordinary conception of sound as *the* things we hear.

We can now easily see the error in this line of thought. An auditory experience of an object represents that object auditorily. It thus has phenomenal properties appropriate to the auditory manner of representation. Let us call such phenomenal properties "auditory phenomenal properties". Similarly, a visual experience has visual phenomenal properties, a tactile experience has tactile phenomenal properties. The error is then that of concluding from the difference in phenomenal properties to the difference in the physical properties represented by the experiences.

Consider the case of shape. We can see and feel the shape of an object. Our visual and tactile experiences of that shape have visual and tactile phenomenal properties respectively, but they do not, therefore, represent visual and tactile shapes. They represent the very same shape but only in different manners.

Accordingly, the phenomenal pitch, loudness, and timbre of an auditory experience are its auditory phenomenal properties. They are, of course, phenomenally different from visual or tactile phenomenal properties. Nonetheless, all these phenomenal properties may well be representing the same physical properties.

In short, if what I have just said is right, then there is no phenomenological reason to deny that sound sources are directly heard. Sound sources can be perceived in different modalities. For instance, I can hear and see a coach's passing-by or hear and feel the vibration of a tuning fork in my hand. The resulting auditory, visual, and tactile experiences have different phenomenal properties, but this does not imply that they represent auditory, visual, and tactile objects respectively. Rather, they are simply auditory, visual, and tactile representations of the same sound sources. It is, therefore, a mistake to conclude based on the auditory phenomenal character of auditory experiences that sound sources are not direct auditory objects.

Young (2016, p. 30, n. 5) nicely call this kind of mistake "correspondence visuocentrism". It is the mistake of "judging whether or not a non-visual modality represents a given environmental aspect on the basis of how similar it is to vision." There is no reason to expect that the same aspect to be represented in the same way across different modalities (ibid., p. 30), not to mention that the visual representation should not be privileged over other modalities in having exclusive access to material objects and events.

There is, therefore, no reason to deny that sound sources can be direct auditory objects. Moreover, admitting that they can be directly represented in auditory experiences provide a simpler account of the auditory phenomenology. Why are direct auditory objects appear to be distally located at where sound sources in fact are? Because they are sound sources themselves, and hence they are simply represented as located at their own locations. There is no need to explain our access to the locations of sound sources in terms of a further account of how hearing sounds can afford such an awareness. The simplicity of a theory of auditory perception which allows sound sources to be direct auditory objects makes it a better theory than those which commit to sonicism.

2.3 Conclusion

In this chapter, I considered sonicism as the contemporary incarnation of the veil of sound. I examined the notion of auditory objects and proposed my definition to guide our further investigation. Two formulations of sonicism are discussed. Although I could not provide any decisive argument against sonicism, I at least provided some strong reasons to reject its status as a theoretical background for theories of sound. Sonicism might be true under a certain theory of sound. It just has no methodologically acceptable role to play in theorising about the nature of sounds.

Close to the end of my discussion, I proposed a diagnosis of why the direct-indirect distinction might seem plausible. From this diagnosis, I drew out how we come to have the ordinary conception of sound as the auditory entities represented by our auditory experiences. I also showed that it is reasonable to accept sound sources as direct auditory objects. This highlights the identification problem of auditory objects, as it makes it more plausible that the direct object of an auditory experience may be a sound source rather than a sound.

Before closing this chapter on auditory perception, I would like to say a little bit more on sounds. Philosophers of sound would probably be dissatisfied with what I said near the end of the chapter. Even if I am right, my speculative story concerns the ordinary concept of sound only. However, the philosophy of sound studies sounds rather than our concept. Therefore, my story is simply irrelevant to their theories.

I disagree. People who fall prey to the mistake of correspondence visuocentrism may wrongly think that what they hear are not sound sources but some merely auditory things. Together with the thought that sounds are such merely auditory things, these people would take their auditory experiences of sound sources as auditory experiences of sounds instead. If I am right that this is a widespread problem, then many perceptual reports of hearing sounds as having such-and-such properties should be re-interpreted as attributing those properties to the experienced sound sources. Philosophers of sound should then be wary of this problem in determining what features can be possessed by sounds. There is then a need to re-examine the features of sounds commonly accepted in the literature to see if they are still acceptable. This is the task of Chapter 3. It will further be seen in Chapters 4 and 5 that many theories of sound are based on the mistake of correspondence visuocentrism and hence should be rejected. The problems in the ordinary concept of sound are relevant to the philosophy of sound because it has been misleading the field into a wrong direction for too long.

3 Features of Sounds

Philosophers often judge the various theories of sound in terms of their compatibility with certain key features of sounds. This strategy might look indisputable: a theory is likely false if it identifies sounds with entities which cannot possess at least some key features of sounds. Even if we want to be more cautious by allowing that our auditory system might systematically misrepresent sounds in some aspects, it is still reasonable to expect that a theory which identifies sounds with entities with more of those key features is more likely to be true. At this general level, I have no issue with this strategy.

My worry concerns the more specific aspect of how this strategy is commonly executed. In §2.1.3, I introduced the identification problem of auditory objects. Limiting only to the two options of an event *x* and *the sound of x*, I claimed that we have no idea which one is represented by an auditory experience. In §2.2.2.4, I discussed correspondence visuocentrism and argued that it is plausible that sound sources are direct auditory objects. This chapter examines eight alleged features of sounds and shows that these two problems significantly reduce the number of such features which can be accepted pre-theoretically. As a result, as we will see in the remaining chapters, many arguments in the current literature should be rejected as based on question-begging assumptions about what sounds are.

3.1 Relation to Auditory Perception

It is perhaps the most central idea in our conception of sound that sounds are what we hear. We may, as I think we should, think that we also hear something else, such as sound sources. But still, sounds are among the things we hear. Therefore, in our ordinary conception, sounds are closely tied to auditory perception.

The precise relation between sounds and hearing is not clear. There are some other related ideas which are not as commonly accepted as the basic idea that sounds are objects of hearing. Consider two examples. First, sometimes it is also claimed that sounds are all we can hear. This is the austere view of the auditory world we discussed and rejected in §2.2. Nothing will be added here.

The second example is the Aristotelian idea that sounds are proper objects of hearing (*De Anima*, II.6). By that, it means that sounds can only be experienced auditorily. This is questionable based on what I said in §2.2.2.4. It is fine to say that auditory experiences of sounds have distinctive phenomenal properties, such that they are phenomenally like no experiences in other modalities. This is, however, no evidence that sounds cannot be experienced, say, visually or tactually. We should determine whether sounds can only be heard from what they are rather than the other way round. For instance, if the sound of a tuning fork is just its vibration, then we should accept that sounds can be felt by holding the tuning fork in hand.

Having rejected these two further ideas, we are now back to the basic idea that sounds are among what we hear. This idea might be challenged by what we saw in Chapter 2. First, the identification problem of auditory objects in §2.1.3 denies any ground within an auditory experience for identifying the auditory object as a sound or a sound source. In §2.2.2.3, I introduced the unification problem of explaining how a sound and its source can be directly heard when there seems to be only one individual represented in an auditory experience. Later, in §2.2.2.4, I argued that it is more plausible to accept that sound sources are direct auditory objects. Putting all these considerations together, it seems the idea that sounds are among what we hear is in great trouble.

One possible way out is to deny that sounds are heard. This idea might be supported by the view that auditory perception can be explained without mentioning sounds, and hence there is no reason to suppose that sounds are heard (Young, 2016). This may lead to the extreme position that denies the existence of sounds. A less extreme position would just hold that sounds exist but are inaudible. However, whatever sounds are in this latter view, it is questionable why there is any need to identify them with sounds. Moreover, the question "What are sounds?" is intended to be asking about the nature of the referents of the word "sound", which is a word among the most basic vocabularies in English (similar for other languages). When an ordinary competent user of English utters the word "sound", she normally is talking about something experienced auditorily. Denying that sounds are heard and identifying sounds with some inaudible entities would then simply switch the topic.

Unless there are very strong reasons, we should avoid such an extreme and revisionary view and stick to the idea that sounds are among what we hear. One option on the other extreme of the spectrum is to fall back on the austere view of the auditory world and deny that sound sources are heard. In-between these two extremes, there are still other possibilities. Casati et al. (2013), for example, identify sounds with sound sources, such that there is no difference between hearing sounds and hearing sound sources. Another possibility is to say that we only hear sounds in the causal sense but not in the representational sense. In other words, sounds cause our auditory experiences of sound sources without thereby being represented. There is thus only one individual represented in an auditory experience. Nonetheless, sounds still have a causal role to play in explaining auditory perception, so they cannot be eliminated from our ontology.

The option I take, as we will see in §6.6, is to provide a novel account of what it is to hear sounds and sound sources. The key idea is to treat sounds as the perceptual means by which we hear sound sources. Details of this account will be spelt out in due course. For the current stage, we just need to note that the difference between all these options does not matter, as they all accept that sounds are objects of auditory perception.

3.2 Public Objects

Sounds are normally considered as public objects, i.e. intersubjectively accessible objects (Lowe, 1981, p. 330). Sounds appear in our experiences to be in the external world. People in a concert hall listen to the same music. They discuss the same piece with their friends after the concert. They may sometimes argue about the beauty of the musician's interpretation. This would make no sense if the musical sounds they hear are just their private sensations like pains.

This common-sense construal of the situation is questionable considering what we saw in Chapter 2. If sound sources are direct auditory objects, perhaps what those concertgoers argue about is the sound source, i.e. the musician's playing of her instrument, rather than the sounds produced. Sounds may be more like pains than we expect. After all, there is a noticeable interpersonal difference in the experiences of the audience—there is no reason for selling tickets at different prices if not for the different sounds at different seats. While they hear the same sound source, the sound heard by one person is audible to no other people. Sitting at the back of the auditorium, I cannot hear the sound people in the front row hear. Even for the person sitting next to me, the sound she hears is nonetheless different from the one I hear. Or so it may be argued.

A likely and reasonable response is that while it is true that people in different seats cannot share exactly the same auditory experience, the difference in their experiences can be adequately explained by their different perspectives (ignore for the moment the individual differences in the auditory apparatus).

There seems not to be any ground to assert that sounds are public objects which does not beg the question against subjectivism of sound. Nonetheless, subjectivism treats sounds as sensations. This conflicts with our understanding of veridical perceptual experiences as representing their *external* causes. Therefore, insofar as we accept that sounds are objects of auditory perception, we should also hold that they are public entities.

3.3 Temporality

There seems to be an intuitive idea about the relative importance of spatial and temporal features for vision and hearing respectively. While visual objects are often distinguished by their spatial features, auditory objects are instead identified in terms of their temporal features. Physical objects can be seen even without anything happening to them. The visual field is segregated into individual objects in virtue of its spatial layout. In contrast, we identify objects of hearing in terms of their temporal development.

O'Callaghan (2007b, pp. 26-27) goes further in advancing the idea that sounds are event-like particulars on this basis. Sounds develop in time. We do not hear a sound in its entirety at a single moment. In contrast, when we see a physical object, it appears to be complete. The difference that physical objects are wholly present while sounds can only be heard in parts suggests that these two categories of entities have different ways of persistence: objects endure but sounds perdure. Sounds are thus similar to events—they last for a while, and only a temporal part of the larger whole is present at each moment.

We shall not take O'Callaghan's argument in its full force. We might agree that there are event-like auditory objects, but he might have misidentified such objects as sounds. If there are other auditory objects which are object-like, then it is at least theoretically possible that sounds are objects instead of events. We should, therefore, examine more carefully if our auditory world is populated by both event-like and object-like entities. I shall show that both our visual and auditory worlds are populated by event-like and object-like entities in roughly the same way. Their apparent difference is just a result of an inappropriately drawn analogy.

A usual way to analogise hearing with vision is to treat sounds and sound sources as the auditory analogues of colours and reflective objects respectively. Even at this general level, it is already a clear disanalogy. While sound sources are the sources of the compression waves reaching our ears, reflective objects merely reflect pre-existing light. Given that we can see light sources, it should be more appropriate to compare them with sound sources.[14]

Sound sources can only be heard when they are sounding. This is not something special about our auditory world. It is also the case that light sources can only be seen when they are emitting light. Notice that I do not deny that a light bulb can be seen when it is off. However, in that case, it is seen in virtue of reflecting light from other light sources. Therefore, it is not seen *qua* light source.

Similarly, it might be possible that a loudspeaker can be heard when it is not making any sound. With sufficient training, blind people can have an impressive ability of echolocation. They might be said to hear the loudspeaker in virtue of the compression waves it reflects. I will discuss echolocation further in §6.9.1. At this stage, we simply need to notice that objects can be heard in much the same way as reflective objects are seen. Objects might not be recognisable through echolocation because of the limitations of our auditory system, but this is no reason to deny that they can be heard in this way. Therefore, it is a mistake to think that our auditory world is only populated by event-like entities. Echolocation involves auditory experiences which unfold and have a dynamic character, but the objects represented can be as static as, say, the words you are now reading.

If echolocation does not seem to be a good example to you, let us consider the case of auditory constancy. When a car is driving toward me, the phenomenology of my auditory experience cannot be fully described in terms of the changing loudness and timbre. There is also a constant element: my experience is and appears to be about some approaching object with constant auditory properties instead of a stationary object with changing auditory qualities. My auditory experience does not

[14] In §6.1, I discuss the analogy between sound and colour and that between sound and light in greater detail.

merely represent the event of driving toward me but also the car. There are thus both an event-like entity and an object-like entity in the same experience.

Here an opponent might say that hearing and vision are nonetheless different. While we see an event by seeing the objects involved, we hear an object by hearing an event involving it. This idea indeed sounds mysterious to me. It seems to be saying that we can hear an event before hearing the objects involved, and those objects are heard only after that.[15] Does that mean that when we first hear the event, it appears as object-less?

The opponent would probably reply that her claim does not concern temporal priority but metaphysical dependence. While event-seeing depends on object-seeing but not the other way round, the converse is true for event-hearing and object-hearing. Nudds (2014a, p. 477) provides a forceful counterexample to the visual part of this claim. Consider the following experiment described by Johansson (1975). In a dark room, if a person has small lights attached to her arm and leg joints, we can only see a cluster of lights if she stands still. However, once she starts moving, we almost immediately see the human figure. Even her gender can be easily identified. If there is any priority between object-seeing and event-seeing, in this case the object is seen in virtue of seeing an event. It is the seeing of the event of the movements of the lights which enables us to see the human figure.

We see and hear objects. Also, we see and hear events. Do we have any second-order experience of the alleged dependence between experiences of objects and experiences of events? I can find no such experience. What seems to be closer to the truth is that, when an event occurs, I experience it as "some object x is φ-ing". This is a single experience which constitutes two facts: the fact that this is an experience of the object x, and the fact that this is an experience of the event of φ-ing. I might characterise this experience as an experience of the object x or an experience of an φ-ing event, but a characterisation reflects what the experience is only indirectly through how we think of it. My attention may be captured by the object x or the event of φ-ing. Sometimes, I may not recognise either of them. There can be many reasons for me to characterise the experience only by reference to one of the two represented entities. Nonetheless, this does not show that there is a different experience corresponding to each characterisation.

---

[15] Young (2018, p. 2938) expresses basically the same worry.

I have tried to describe how a better analogy between vision and hearing can be drawn. If it works, then we should see that our auditory world, as well as our visual world, are populated by both object-like and event-like entities. More precisely, what our perceptual experiences represent are complex entities which are structured in the form "*x* is *φ*-ing". The identification problem of auditory objects forbids us to say which entities should be identified with sounds, and hence there is no neutral ground for saying that sounds are event-like entities. The temporality of our auditory experiences might have no direct bearing on what sounds are.

3.4 Spatiality

An ordinary auditory experience represents something as located at a certain distance away from the perceiver. Normally, that location is the location of the sound source. If you close your eyes and reach out to that place, you will be able to touch the sound source. Interestingly, some philosophers share the general opinion that the thing which auditorily appears at that location is a sound rather than the sound source. Accordingly, assuming the veridicality of the auditory experience, they generalise the case and conclude that sounds are located at or near their sources.

Our reflection in Chapter 2 shows that such a conclusion is unwarranted because of the identification problem of auditory objects. It is indisputable that auditory experiences do represent certain entities at the locations of the sound sources. However, the distally located objects may well be the sound sources themselves. If so, are sounds represented as located?

One possibility is that sounds are represented as located at the same locations with their sources. This means both sounds and sound sources are represented at the same locations. It is unclear how exactly this could be the case. As I pointed out in §2.2.2.3, when an audible event happens, there appears to be only one individual in my auditory experience. Those who opt for this possibility would then need to solve the unification problem. Such a solution would be available after a theory of sound is provided. At our current stage, the problem is that there is no clear reason why we should assert that there is anything at the locations of sound sources in addition to the sound sources themselves. Why should we split the individuals represented in our auditory experiences into two entities? There is no phenomenological ground for unnecessarily multiplying the number of entities experienced.

Another possibility is that sounds are represented but not spatially. Even if sounds have determinate locations, it is not *a priori* necessary that their locations should be represented in auditory experiences. If, for example, our auditory system has no way to access the locations of sounds, then, of course, we do not experience sounds as located. To determine if this is the case, we need to know how our auditory system works. Moreover, the nature of sounds would also limit how their locations could be accessed by our auditory system.

The general lesson is that there are multiple ways to characterise the spatial phenomenology of auditory experiences, and any characterisation which would be rich enough to indicate how sounds are represented spatially would inevitably have implicitly assumed a certain theory of sound and a certain theory of auditory perception. As a result, instead of using any of such characterisation to argue about the nature of sounds, we should choose the best characterisation based on the theory of sound and theory of auditory perception which are accepted on an independent ground. For now, we can at most say that granted that sounds are represented in auditory experiences, in the simple scenario where there are only a sound source and a perceiver, sounds are either not represented spatially or represented at or near their sources, as the only location represented in the auditory experience is the location of the sound source. More locations may be represented in more complex situations. For example, we may include the locations of reflective surfaces if echoes are heard. Nevertheless, the general point still holds: the represented locations are all occupied by entities which need not be identified with sounds, such that saying that sounds are there may unnecessarily multiply the number of auditory objects.

3.5 Relation to Ordinary Material Objects

It is sometimes suggested that sounds are distinct from ordinary objects which we can see around us. Furthermore, it is also said that sounds are independent from those objects. I shall examine these two claims in turn.

The first claim is held by Nudds (2001, p. 210). He takes the distinction between sounds and ordinary objects as so compelling that "[a]ll you have to do to confirm it is close your eyes and reflect on the character of your auditory experience." Similarly, O'Callaghan (2007b, p. 5) claims that sounds do not appear to have properties such as shape, size, mass, solidity, etc. which are typical for ordinary objects. Both authors seem to take the distinction as obvious and indisputable, but I cannot see how it can be.

To begin with, notice that both Nudds and O'Callaghan do not talk about sounds and ordinary objects directly but only through their appearances in auditory experiences. This is understandable, as we do not know what sounds are and hence there is no theoretically neutral ground for distinguishing them from ordinary objects. However, the appeal to auditory experiences creates its own problem.

When an audible event happens, there appears to be only one individual in my auditory experience. If it is a sound, then the ordinary object undergoing that audible event does not have a distinct appearance which can be compared with the appearance of the sound. The same is true if that individual is the audible event. If that individual is a unified representation of the sound and the audible event, then they simply have no distinct appearances. Therefore, if Nudds is right that the distinction between sounds and ordinary objects is phenomenally evident, this distinction cannot be drawn by comparing distinct appearances.

As a result, Nudds would need to say either that ordinary objects auditorily appear to have features which cannot be attributed to sounds, or that sounds auditorily appear to have features which cannot be attributed to ordinary objects. The first option can be rejected, as we do not know what features of ordinary objects cannot be attributed to sounds before knowing what sounds are. As for the second option, the most likely proposal would say that those features are auditory properties including pitch, loudness, and timbre. However, we have already seen in §2.2.2.1 that auditory properties can be attributed to event sources. It is then possible to attribute auditory properties to ordinary objects derivatively when they are undergoing some audible events. Apart from auditory properties, it is not clear what other properties of sound cannot be attributed to ordinary objects.

O'Callaghan's proposal is of no help here. From the claim that sounds do not appear to have certain properties, it does not follow that they do not have those properties. At most, we can say that those properties, when possessed by sounds, are inaudible. Appealing to the veridicality requirement would not help, as the claim is not that sounds appear not to have shape, size, mass, solidity, etc. but that sounds do not appear to have those properties. An experience representing sounds as not having properties they do have is illusory, but an experience not representing some properties of sounds is limited but can still be veridical.

On the one hand, there is no theoretically neutral ground for distinguishing sounds from ordinary material objects. On the other hand, we cannot see how this distinction can be supported by auditory experiences. We should therefore not assume this distinction in theorising about the nature of sounds.

As for the claim that sounds are independent from ordinary objects, it is also held by O'Callaghan (2007b, p. 6):

> Sounds as we experience them in hearing are audibly independent from ordinary material objects in a way that colors and shapes are not visibly independent from objects. The sounds seem produced or generated by ordinary objects and events; colors and shapes do not.

It should be clarified what sense of independence O'Callaghan is talking about, as being produced or generated by sound sources could be understood as showing that sounds are *causally* dependent on sound sources. From the contrast with colours and shapes, it seems O'Callaghan might be saying that auditory experiences represent sounds as being detached from their sources, while visual experiences represent colours and shapes as attached to their bearers.

O'Callaghan's claim here concerns the phenomenology of auditory experiences, so it can be significant for the question about the objective nature of sounds only on the assumption that auditory perception is veridical in this respect. We might grant this assumption for the moment. What we should ask is how auditory experiences could represent the alleged detachment of sounds from their sources. In an ordinary auditory experience, are there separable auditory objects which correspond to a sound and its source respectively? I have already rejected this object-splitting view of auditory experiences. What we have instead is an individual auditory object which is a unified whole, and we have no clue as to whether it is a sound or a sound source. The unity might just be an appearance, but as far as phenomenology goes, there is no ground to say that any alleged detachment can be experienced. Therefore, we also should not accept the claim of independence.

3.6 Being Caused by Sound Sources

We have just examined O'Callaghan's claim that sounds are represented as being produced by and independent from their sources. Although we have rejected this claim, it seems our very concept of sound sources trivially implies that they are objectively causes of sounds. Moreover, we may agree in most scenarios what the sound sources are. It thus seems to be indisputable that sounds should be entities caused by those things which we identify as sound sources regardless of how our auditory experiences represented them and their relations.

However, do we experience the causing of sounds? Nudds (2001, p. 220) argues that we do, but only in virtue of bi-modal experiences. We hear a sound, see or touch its source, and experience the causation between them. In general, it is hardly deniable that we can experience causation bi-modally. I press the light switch while looking only at the light across the room, and then see the light turns on. I bi-modally experience my pressing the switch causes the turning on of the light. I kick a stone into an abyss which is too deep for its bottom to be seen. A few seconds later, I hear the stone hits the bottom. I bi-modally experience the causation between my kicking the stone and my hearing it hits the bottom. If these cases are what Nudds is talking about, then we shall happily concur with him.

Unfortunately, the cases mentioned in his paper are different. The two cases we just saw involve non-simultaneous causes and effects. In contrast, the causes and effects of Nudds's examples are all simultaneous. Let us focus on one of his examples—a tuning fork (ibid, p. 218). According to Nudds, the cause is the vibration of a tuning fork, while the effect is a sound of the tuning fork. We feel the vibration, hear the sound, and experience the sound as caused by the vibration.

Although Nudds's description comes close to how ordinary people would think about the case, I do not think that he describes it correctly. Nudds's description requires that the vibration and the sound are distinct entities which are experienced to be causally connected. However, I have argued that event sources can be represented directly in auditory experiences. Also, the distinction between an event source and a sound of that event source is not phenomenological. It seems to be more plausible that what Nudds refers to as the sound *is* the vibration: there is no cause and effect, but only one and the same vibration experienced both by touch and audition simultaneously. Bi-modal experiences of sound sources like those mentioned by Nudds represent no causal structure. Therefore, the phenomenology of those experiences does not tell us anything about the production of sounds.

This gives rise to a new question: How do we acquire the concept of sound source? Is there another kind of experience which can represent the production of sound? Or do we learn the concept of sound source via non-experiential means? Either way, our possession of this concept suggests that there are some occasions in which we can in some way distinguish between sound sources and their causal products. The fact that the concept of sound source is pre-scientific also suggests that such occasions should be fairly common in ordinary life. I will return to this issue in §6.1.2. For the moment, I simply take our possession of the concept of sound source as showing that sounds are caused by sound sources.

3.7 As Bearers of Auditory Properties

Auditory properties should not be confused with audible properties. While all auditory properties are audible, audible properties may not be auditory. If I am right, duration is audible, but it is not auditory (nor is it visual). The three and plausibly the only three auditory properties are pitch, loudness, and timbre, which are closely related to the frequency, amplitude, and spectrum of a stimulus respectively. However, the relations between these two groups of properties are too complicated to warrant any but the most sophisticated reductivist treatment of the three auditory properties. It is therefore tempting to concur with scientists to count the three auditory properties as subjective properties.

Sounds are standardly considered as bearers of auditory properties (Nudds, 2010a, p. 106; O'Callaghan, 2007b, p. 17). If auditory properties are subjective, then we might need to identify sounds with some subjective entities. One simple way to avoid this anti-realist consequence is to take auditory properties as response-dependent. This implies that sounds must be entities which can be represented by auditory experiences. Otherwise, they could not possess any response-dependent properties and hence could not be bearers of auditory properties so conceived.

This implication fits nicely with the generally accepted idea that sounds are auditory objects. However, we shall proceed more carefully. Given the identification problem of auditory objects, we may only say that auditory objects are represented as bearers of auditory properties without identifying them as sounds. It seems then that we can accept at this stage the claim that sounds are bearers of auditory properties only if all auditory objects bear auditory properties.

Unfortunately, there seems to be auditory objects which do not have any auditory properties. Consider hearing the traffic outside the window. We can hear the sound as being blocked by the window, or we hear the sound as being blocked by some barrier (it might be controversial whether the window can be recognised as such in the experience). The auditory awareness of the muffled sound as resulting from the presence of some barrier implies that the barrier is represented in the experience. The barrier, i.e. the window, however, is not represented as having any pitch, loudness, or timbre.

If it is accepted that barriers can be represented in auditory experiences, then at least some auditory objects do not have auditory properties. That there is at least one possible way for an auditory object to have no auditory properties gives us a reason to expect that there might be some other possible way. It seems then we might need to address the possibility that sounds might be represented in auditory experiences but not as having any auditory properties.

A likely objection here would be that such a possibility is too far-fetched to be taken seriously. It seems sounds are the only auditory objects which can have auditory properties. So, even if there are other auditory objects which have no auditory properties, it simply does not matter. A possible response is to show how auditory properties can be attributed to auditory objects other than sounds. In §2.2.2.1, I have already argued that auditory properties can be attributed to event sources. Another proposal can be found in Young (2016, pp. 81-89), where loudness is reattributed to the force applied to a thing source, pitch is reattributed to the thing source, and timbre is "dissolved" as merely a way we describe the overall character of an auditory experience.

If we accept any of these proposals, then two possibilities follow. Perhaps auditory properties can be attributed to sounds *and* other entities such as sound sources, or perhaps auditory properties should be *re-attributed* to non-sound entities, such that sounds bear no auditory properties. Consider the second possibility first. If sounds have no auditory properties, then what other properties can they be represented as having? Their locations can only be represented if they are located at their sources. How about material properties like size, mass, solidity, etc.? These properties are possessed by sound sources which also possess auditory properties under the current proposal. If sounds are represented as having material properties but not auditory properties, it seems they should be some material objects other than the sound sources. However, in typical scenarios, the only entities capable of having those material properties in the surrounding of the hearing subjects are the sound sources. Assuming the veridicality of typical auditory experiences, it implies the inconsistent result that sounds are and are not sound sources, so sounds should not be represented as having those material properties.

It might then seem that the only remaining candidate is some temporal property. However, in the absence of other properties, sounds cannot be represented as changing or remaining unchanged. It is thus inconceivable how they could be represented as having any temporal property.

It seems then sounds could only be represented as bare particulars, but it is highly implausible that bare particulars appear in perceptual experiences. To conclude, under the proposal that auditory properties should be reattributed to sound sources, it seems sounds cannot be represented in auditory experiences. This contradicts with our overarching assumption that sounds are auditory objects. We should then reject the second possibility and accept that auditory properties can be attributed to sounds and non-sound entities such as sound sources.

It seems sound sources can possess auditory properties only derivatively. It is because the sounds they produce have certain auditory properties that they possess the corresponding auditory properties. In other words, sounds are the *primary* bearers of auditory properties. The piccolo is high-pitched because the sounds it produces are high-pitched. The trumpet is loud because its sounds are loud. The violin is bright because, in general, its sounds, especially those on the E-string, are bright. We may need to separate the case of thing sources from the case of event sources. As we saw in §2.2.2.1, objects are sound sources in virtue of undergoing some audible events. Presumably, it seems the attribution of auditory properties to thing sources is parasitic on the attribution of auditory properties to event sources.

The last example of violin brings out the further condition that an auditory property can be attributed to a sound source without it being the case that every sound produced by that source has that property. However, it is not clear what the best account of such attribution should be. It might be determined by the proportion of sounds which have the attributed property. Or the statement which attributes the auditory property to the sound source should be understood as a generic statement. Or there might still be other possibilities. Since this issue is tangential to my discussion, I shall leave it open here.

In sum, the commonplace idea that sounds are bearers of auditory properties survives our critical assessment. Therefore, the best theory of sound should identify sounds with entities which can possess all the three auditory properties.

3.8 Survivalism

An idea closely related to the claim that sounds are bearers of auditory properties is that sounds survive changes in auditory properties (O'Callaghan, 2010a, p. 250). This view can be called "survivalism" (Cohen, 2010). It is also claimed that the identities of sounds hinge on how their auditory properties change over time (Nudds, 2015, p. 275). For example, the wail of a siren and the bark of a dog are different sounds because of the very different ways their auditory properties change respectively. That a sound can maintain its identity through changes in its auditory properties is sometimes treated as evidence for the event-like nature of sounds.

Although I accept that sounds are events because they are compression waves and compression waves are events, I do not accept survivalism as a theoretically neutral ground for theories of sound. More precisely, the phenomena which are employed to support survivalism can also be described as the sound sources survive changes in auditory properties without implying that their sounds survive changes in auditory properties or not. Therefore, unless sounds are identified with sound sources, those phenomena do not support survivalism, and hence we need not accept it. The truth of survivalism can be ascertained only after we know what sounds are, so it is question-begging to ground any theory of sound in it.

Consider the case of a piano. When someone presses a key on a piano, a hammer strikes a string (or a few strings, depending on which note is played). The conventional way to describe what happens is to say that a sound with a sharp attack is produced, and then it decays gradually. In other words, a sound undergoes a characteristic change in its loudness, and hence survivalism is assumed.

Once we allow that sound sources can be directly represented in auditory experiences, we can easily see an alternative understanding of the situation. In short, that entity which appears to change in loudness is the piano itself rather than a sound as a distinct entity. There can be room for further clarification. For instance, loudness might be treated as an object property attributed to the piano, such that what happens is the piano changes its loudness. Alternatively, loudness might instead be treated as an event property attributed to the oscillation of the string(s) and the soundboard of the piano. We do not need to engage in this further debate. What we need is simply the claim that the individual which survives the change in loudness is the thing source or the event source rather than a sound.

Intuition might favour the conventional picture. However, a consideration of a visual analogy should give us reason to suspect the reliability of intuitive judgements. Imagine a mysterious metal called "hammerluminium". Hammerluminium, as its name suggests, has a very peculiar property that if someone hits it with a hammer, it glows, and the brightness reduces gradually. Also imagine a hammering machine which works just like a piano, with the only difference being that the strings are all replaced with pieces of hammerluminium. When someone presses a key on the hammering machine, one piece of hammerluminium glows and then dims gradually. There is an entity which survives the change in brightness. Is it the piece of hammerluminium or some distinct entity?

Following the conventional way of conceiving the case of a piano, it seems we should say that the entity which survives the change in brightness is something distinct from the piece of hammerluminium. However, to my intuition, it is more natural to say that it is the piece of hammerluminium which changes in brightness. At least, this understanding is equally, if not more, sensible than the former one.

The analogy between the hammering machine and the piano makes it a reasonable option to understand them in the same way. If it is sensible to say that the piece of hammerluminium survives the change in brightness, then it should be equally sensible to say that the piano (or the oscillation of its string(s) and soundboard) survives the change in loudness.

It seems sonicism is a major reason for accepting the conventional description in the case of sound. If sounds are our direct auditory objects, and direct auditory objects appear to survive changes in their auditory properties, then it follows that at least in veridical experiences the experienced sounds survive changes in their auditory properties. Once we are liberated from the constraint of sonicism, we can easily see that the conventional description is guilty of unnecessary complication. It is indisputable that the amplitude of the oscillation event of the string(s) and the soundboard of the piano changes, and the piano as the thing source changes in the sense of undergoing a changing event. However, the phenomenon is neutral to the truth of survivalism. Perhaps what is produced by the sound source is a sound with temporal parts of different loudness, or it is rather a stream of multiple sounds each of which instantiates different loudness. These are "alternative and equally acceptable methods of bookkeeping" (Cohen, 2010, p. 312). Nothing about sounds specifically can be learnt from such phenomena.

A clarification is needed here. I am not denying that survivalism might be true. My challenge is merely that there is no non-question-begging description of the relevant phenomena. The same phenomenon can be described by both survivalist and non-survivalist descriptions. Without first knowing the nature of sounds, we do not know which description is right. Therefore, the relevant experiences cannot ground any theory of sound.

If survivalism is false, then it is wrong to say that sounds can be identified by their characteristic changes in their auditory properties. Each sound has a unique combination of pitch, loudness, and timbre. It exists if there is no change to its auditory properties and is replaced by another sound if there is any such change. What we should say instead is that thing sources or event sources can be identified by their characteristic sequences of sounds of different auditory properties. They might then be said to survive changes in auditory properties derivatively.

3.9 Summary

In this chapter, we have examined eight alleged features of sounds. They are often employed as evidence for various theories of sound. However, we can now see that some of them depend on the problematic assumptions exposed in Chapter 2. As a result, only half of them can provide a neutral starting point for further investigation into the metaphysical nature of sounds.

We accept that sounds are objects of auditory perception. To this extent, they are public entities. It is trivially true that sounds are caused by sound sources. Lastly, although sounds are not the only bearers of auditory properties, they are the *primary* bearers of those properties. Everything else can only bear auditory properties derivatively.

In contrast, I argued that auditory experiences do not represent sounds as distinctively temporal in comparison to perceptual objects in general. I also argued that sounds may not be spatially, but if they were, they would be represented as located at the locations of their sources. As for the claim that sounds are distinct or independent from ordinary material objects and survivalism, they are not supported by the phenomenology of auditory experiences.

The general lesson of this chapter is that the metaphysics of sound should not rely too much on the phenomenology of auditory experiences. There can be different ways of characterising the phenomenology, and they may have already assumed certain things about the nature of sounds. If it is possible, we should appeal more to objective, non-experiential evidence. Auditory perception can still be relevant, but we should examine its objective conditions rather than focus on the phenomenology of auditory experiences. After all, how we experience sounds to be is constrained by the nature of sounds but not the other way round, and we do not seem to know the phenomenology of auditory experiences well enough.

4 Theories of Sound (I)

Having now examined the commonly accepted features of sound, we are now at a better position to evaluate the various theories of sound, especially in terms of how well they are grounded in the phenomenology of auditory experiences. In this and the next chapters, I move on to examine the existing theories of sound in the philosophical literature. Despite the short history of contemporary philosophy of sound, many theories have been proposed. In order to highlight the differences and similarities between them, §4.1 presents a taxonomy of these theories. Five of them are criticised in §4.2. The rest are the topic of the next chapter.

4.1 Taxonomy

There are two common ways to classify theories of sound—in terms of the ontological categories of sounds or in terms of the location of sounds along the path leading from sound sources to hearing subjects. Main categories for the first scheme include event, property, abstract individual, and sensation. As for the second scheme, adopted by Casati and Dokic (2010) for example, theories are classified into distal, medial, proximal, and a-spatial. Not every theory fits easily into either classification.

Both classification schemes are too simple to reflect the diversity of the existing theories in the literature. To provide a better structure for my discussion in the next section, this section presents a finer-grained taxonomy (Figure 1).

Starting from the top, we quickly come to a theory omitted by the two simple classification schemes—*eliminativism*. This radical position is recently proposed by Young (2016). Young's position admits of two interpretations. The weaker and more charitable one merely says that we do not hear sounds, but this is compatible with the objective existence of sounds. This interpretation fits better with Young's overall focus on auditory perception in his work. In contrast, the stronger interpretation holds that the existence of sounds should be rejected. This stronger interpretation is suggested at a place where Young says that throwing out sounds can avoid the metaphysical question concerning whether sounds and sound sources are

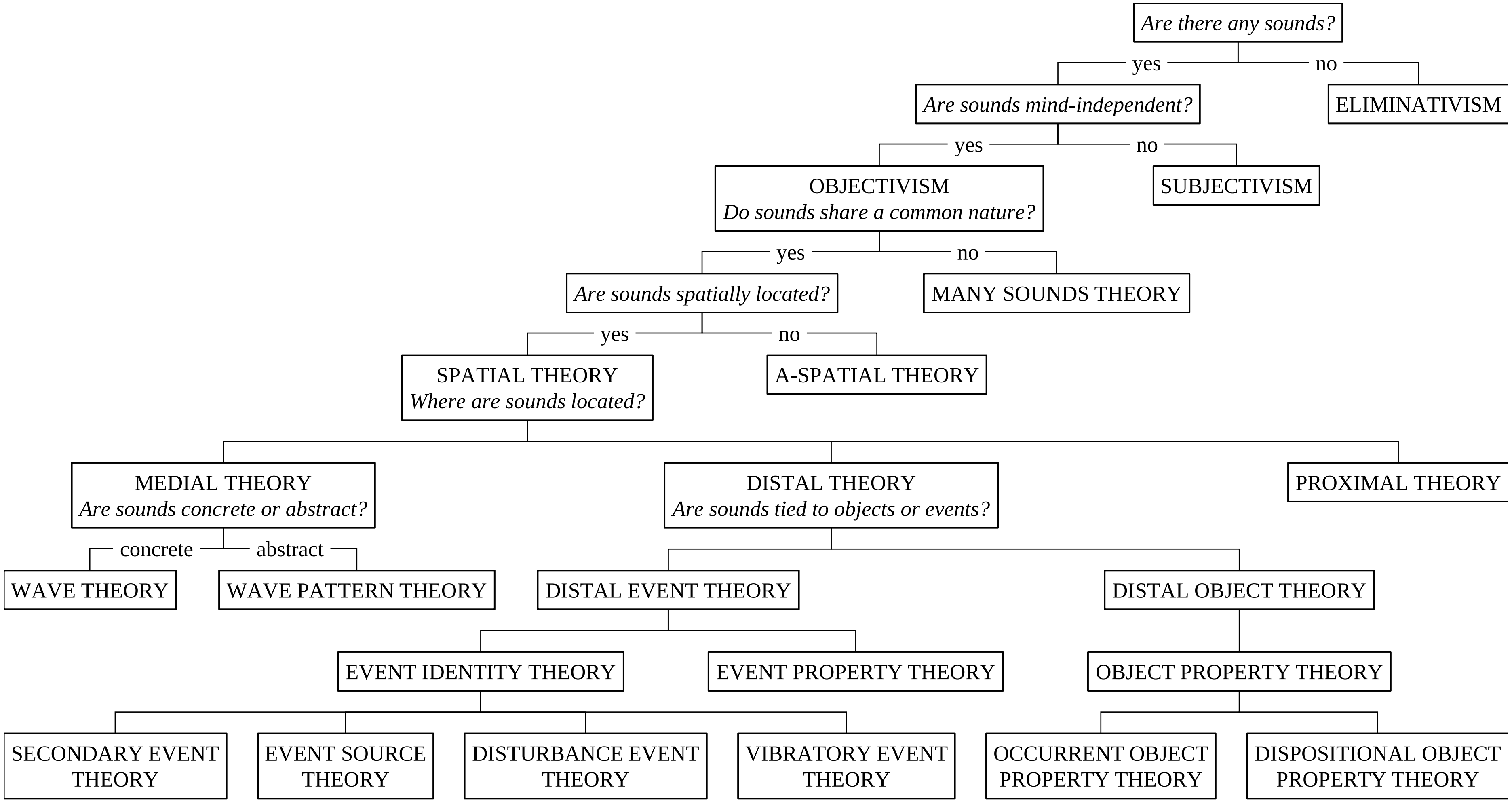

Figure 1. Taxonomy of Theories of Sound

mereologically or causally connected (ibid, pp. 17-18). Merely denying the role of sounds in hearing without also rejecting the existence of sounds cannot warrant such a dismissive response to this metaphysical question. For the sake of my discussion on sounds, I assume the stronger interpretation, and hence count Young as an eliminativist.

Young's main reason for eliminating sounds from our ontology is that auditory perception can be fully accounted for without incurring any sound. He takes our auditory world as exhausted by thing sources, event sources, and empty spaces. He shows that our auditory experiences of each of these categories can be explained without supposing further that we hear sounds. Since sounds are not themselves objects of auditory experiences, nor do they play any intermediary role in auditory perception, they should be rejected for the sake of parsimony.

For those who are not attracted to this extreme view, the next dividing line concerns whether sounds can exist mind-independently. *Subjectivism* includes those views which take sounds as mental entities such as sensations or ideas. Maclachlan (1989, p. 30; 2013, p. 6) is probably the only contemporary philosopher who claims that sounds are sensations, but of course, subjectivism can be traced back to early modern empiricists like Berkeley (1713/1979, p. 19) and Hume (1739-1740/2007, p. 301). Some scientists claim that the word "sound" refers to either a compression wave or a sensation (e.g. Rossing, 2014, p. 1). Subjectivism can to that extent be considered as embraced by some members of the scientific community.

In contrast, the majority view in contemporary philosophy of sound is *objectivism*. Most theories of sound suppose that sounds share a common nature. This is, however, not the only possible option. What I call "*many sounds theory*" is the view that sounds form a heterogeneous kind. One possible way to cash out this theory is to say that the only reason why members of such sub-categories are all sounds is that all of them can in principle be our auditory objects. Such sub-categories might also include entities inaudible for us because of the limitations of our auditory system. Nonetheless, if our auditory system were more sensitive to the kinds of property we can in fact hear, such entities would then become audible. The existence of such a possibility is sufficient for counting such entities as sounds.

This unusual position has been proposed only very recently. In connection with his discussion on the perspectival aspect of auditory perception, Kulvicki (2017, pp. 90-91) proposes two possible views of sound. The first one identifies sounds with perspectival features which are understood to be abstractions over intrinsic and relational features of ordinary objects, events, and environs. The second and more catholic proposal takes sounds as comprising such perspectival features as well as intrinsic features of ordinary objects and happenings. This latter view can be counted as a many sounds theory, and it is the one favoured by Kulvicki.

As for Kulvicki's first proposed view, its identification of sounds with abstractions invite the question "Are sounds spatially located?" This question draws the next dividing line in our taxonomy. *A-spatial theory* denies the spatiality of sounds. The most well-known proponent is Strawson (1959, pp. 65-66), and now we may also group Kulvicki's first view into this camp, as abstractions are not in space and time.

Spatial theory, which locates sounds in space, can be divided into three groups using the second classification scheme introduced at the beginning of this section. These three groups of theories locate sounds respectively at the proximal stimuli, at the distal stimuli, and in the medium. The first one is *proximal theory*, which ties sounds to proximal stimuli. So defined, it has no advocate in the literature. Casati and Dokic (2010) treat subjectivism as a proximal theory, but I reject this classification, as it is preferable to stick to the common usage of the word "proximal" in the psychology literature to refer to proximal stimuli. Sensations are not proximal stimuli—rather, they are products of proximal stimuli.

The next group of spatial theory is *distal theory*, which is the majority view in the contemporary debate. Distal theory locates sounds at or near sound sources. Depending on whether sounds are tied to objects or events, distal theory can be divided into *distal object theory* and *distal event theory*. *Occurrent object property theory*, the distal object theory proposed by Pasnau (1999, p. 316), claims that sounds either are or supervene on the vibrations of thing sources. Those vibrations are treated as occurrent properties of thing sources.

Another theory which also identifies sounds with properties of thing sources is Kulvicki's earlier view. He defends in his (2008a, p. 2) and (2015, p. 206) the stable disposition view of sound. According to this *dispositional object property theory*, sounds are dispositions of thing sources to vibrate when stimulated. When someone strikes a drumhead, a sound is revealed instead of produced.

Distal event theory, the more popular version of distal theory, instead ties sounds to events. There are two main variants: *event identity theory* which identifies sounds directly with events, and *event property theory* which treat sounds as properties of events.

There are four kinds of event identity theory in the literature. The first one is *vibratory event theory*. This theory was first proposed by Casati and Dokic (1994) in French, which remains largely unknown in the English-speaking world. These two authors later presented their view in English in their (2009) and (2010, §3.2). Matthiessen (2010, pp. 94-96) also advocates this view. Like Pasnau's occurrent object property theory, this view also identifies sounds with vibrations of thing sources. There are two main differences though. First, while vibratory event theory treats vibrations of thing sources as events, Pasnau treats them as properties. Second, Pasnau, but not vibratory event theorists, allows the alternative that sounds merely supervene on vibrations.

A thing source can vibrate in a vacuum, but no sound can be heard. Vibratory event theory implies that there is an inaudible sound. The existence of an elastic medium is required for merely the audibility but not the existence of a sound. However, it can also be argued that there is simply no sound at all in a vacuum, such that the existence of a sound depends on the presence of an elastic medium. *Disturbance event theory*, defended by O'Callaghan (2007b, 2009c), captures this alternative view by saying that a sound is the event of its thing source's disturbance of the surrounding medium. Such a disturbance event is a relational event which happens between the thing source and the medium, and hence it cannot exist in a vacuum.

Both vibratory event theory and disturbance event theory distinguish the events identified as sounds from their event sources. Casati et al. (2013, p. 463) deny this distinction and directly identifies sounds with event sources. I call this view "*event source theory*". Hearing a bang is just hearing a collision. The three authors simply call their view "the identity view". I reject this label, as it is both uninformative and misleading. It is uninformative because it does not indicate which kind of events are sounds; it is misleading because it might suggest that this theory is special in identifying sounds with events, but the fact is that other kinds of event identity theory also accept such an identity. To the extent that event source theory denies the status of sounds as distinct entities, it may alternatively be counted as a version of eliminativism.

The last event identity theory is *secondary event theory*, proposed by Scruton (2009, 2010). Scruton explains that sounds as secondary events are “secondary” in the sense that how they appear to normal observers is essential to their existence in a way similar to how secondary properties are defined (2009, p. 57). Also, he claims that sounds are pure events which happen without happening to any object (ibid., p. 50). Scruton thinks that these features of sounds are fundamental to the art of music (ibid.).

Proponents of event identity theory also include Matthen (2010, p. 63) and Pasnau (2009, p. 356). However, since they do not make it clear which kind of events they identify with sounds, I do not put them in the four groups above.

Instead of identifying sounds directly with events, event property theory takes sounds to be properties of events. Leddington, the sole proponent of this theory in the literature, rejects the claim that sounds appear to be independent from their sources, a claim which we discussed back in §3.5. He (2014, pp. 330-331) argues that sounds appear to permeate or saturate their event sources similarly to how colours appear to permeate or saturate their bearers. We do not hear sounds and their event sources independently, but instead hear “*noisy events*”. In his (2019, p. 624), Leddington further claims that sounds are event properties *constituted* by the auditory properties of their bearers.

Lastly, we come to the third spatial theory—*medial theory*. All the theories mentioned so far are probably known only within a small circle of philosophers. For the general public, especially those who have received formal education in natural science, they probably only have heard of one theory—*wave theory*. According to wave theory, sounds are compression waves. Apart from scientists, a few philosophers have tried to defend wave theory, e.g. Kalderon (2018a; 2018b, Chapters 3-4); Perkins (1983, pp. 186, 175); Sorensen (2008, pp. 282-284; 2009, pp. 138-139). Some other philosophers simply assume the truth of wave theory, e.g. O’Shaughnessy (2000, p. 445; 2009, p. 117);[16] Tye (1995, pp. 149-150); von Wright (2000).

[16] Strictly speaking, O’Shaughnessy explicitly assumes wave theory only in his (2009, p. 117) but not in his (2000, pp. 445-447). However, the discussion in the latter work shows that he at least thinks that sounds behave in a wave-like manner. For the sake of simplicity, I take both works as assuming wave theory.

It is worth mentioning that Casati and Dokic (2009, p. 103; 2010, §1.2) mistakenly count O’Shaughnessy’s view as identifying sounds with proximal stimuli. Although O’Shaughnessy does

Compression waves are concrete entities in the medium, or at least this is the standard way to conceive of them. Instead of directly identifying sounds with compression waves, another way to locate sounds in the medium is to say that sounds are abstract entities the instances of which are located in the medium. The only example of such a theory in the literature is proposed by Nudds (2009, 2010b). Sounds are patterns or structures of frequency components instantiated by compression waves (2009, p. 75). They are particularized types or abstract individuals like symphonies and words, which are capable of multiple instantiations in space and time (pp. 76-77). Furthermore, since qualitatively different patterns can be instantiated at different locations while both are counted as the same sound, two patterns can be the same sound only if the compression waves instantiating them are produced by the same events (pp. 290-292). I will simply refer to Nudds's view as *wave pattern theory*.

The taxonomy presented here helps us see the unexplored possibilities. For example, event property theory might be further divided into *occurrent event property theory* and *dispositional event property theory*. Arguably, Leddington's theory is an instance of the former. In addition to object property theory, distal object theory might have another variant—*object identity theory*. Moreover, both distal event theory and distal object theory might find their counterparts which identify sounds with *abstract entities* instantiated on event sources or thing sources respectively. These are not all possible additions to the current taxonomy. Anyway, I take it as beyond my purview to show that such unexplored options are serious contenders for the correct theory of sound, unless there is an obvious reason in support of them. I, therefore, omit the unoccupied theoretical space in my taxonomy and move on now to focus on the existing alternative theories in the literature.

4.2 Some Alternative Theories

---

claim that the sound we hear is where we are, that does not mean that there is no sound elsewhere. Interestingly, Casati and Dokic do note that O'Shaughnessy's view "leaves open the possibility that an unheard sound be located away from the hearer" (2010, §1.2). This is in fact not merely a possibility, but is implicit in O'Shaughnessy's discussion, immediately after the passage quoted by Casati and Dokic, about the "spatial and temporal history" of sounds.

I will argue for wave theory in Chapter 6. Before that, I will critically evaluate all the other theories of sound. Chapter 5 will cover all versions of distal theory. Wave pattern theory will be discussed in §6.7. The rest—eliminativism, subjectivism, many sounds theory, a-spatial theory, and proximal theory—are our focus in this section.

4.2.1 Eliminativism

We saw earlier that Young (2016) rejects the existence of sounds for the sake of parsimony, as sounds are neither auditory objects themselves nor perceptual intermediaries. There are two crucial assumptions. First, it is assumed that the only two roles sounds can play are auditory objects and perceptual intermediaries. This assumption would be acceptable if we adopt the weaker interpretation of Young's view, which only denies the role of sounds in auditory perception. As for the stronger interpretation, sounds may well be entities which exist but are not perceived. Even if they are completely irrelevant to our auditory experiences, they may still exist and play some other role in the physical world. Therefore, the first assumption would then be unacceptable. Sure, as I said in §3.1, it would be unclear why there is any need to identify sounds with such inaudible entities. Nonetheless, it is still a possibility which should not be ignored.

The second assumption is that our auditory objects include thing sources, event sources, and empty space only. We may grant that these three categories are all represented in auditory experiences. The problem is that it seems we hear more than only those three. In particular, at least on some occasions, compression waves are also represented. The crucial cases are echo experiences.

Simply put, echoes are compression waves reflected by objects. However, not all echoes elicit echo experiences. Allegedly two features of distinctive echo experiences set them apart from other auditory experiences, and both are related to the length of the delay between the arrival of a direct wave and that of its echo. First, the echo and the direct wave are represented as distinct objects in auditory experiences. This requires that the delay must be long enough; otherwise, an effect called "echo suppression" would take place, meaning that our auditory system would fuse the echo with the direct wave into a single individual. The threshold for the activation of echo suppression varies for different signals. For brief signals like clicks, the upper limit is about 5 milliseconds, but fusion could take place for a delay up to

40 milliseconds in the case of complex signals such as speech or music (Moore, 2013, p. 267). Variations within the limit can change the qualities but not the numbers of the represented objects.

The second feature of distinctive echo experiences is that the distinct auditory objects elicited by a direct wave and its echo are heard to be related. O'Callaghan (2007b, p. 127) points out that if the delay between the echo and the direct wave is larger than 2 seconds, the resulting experience would be of two distinct but unrelated auditory objects.[17]

Echo experiences have been discussed by a few authors in the philosophical literature, but Young does not offer his own account. He does mention echo experiences, but his description is just enough to show that reverberation experiences are different from echo experiences. A positive account of echo experiences according to his sound-less account of audition is missing. He (op. cit., p. 69) seems to concur with O'Callaghan (op. cit., pp. 128-129) in taking echo experiences as misinformative both spatially and temporally about the event sources, but he adds that the approximate locations of reflective surfaces are represented in echo experiences. To this extent, Young is less dismissive of echo experiences than O'Callaghan.[18]

Here we can already notice that Young's official list of auditory objects fails to include reflective surfaces, despite his allowance that they can be represented in echo experiences. This is, however, not the end of the story, as compression waves should also be added to that list. Consider the case of multiple echoes. Imagine a row of buildings separated with appropriate distance, such that when a firework explodes in the nearby area, a listener standing at a certain location can hear a series of reflected compression waves after the direct wave from the explosion. According to Young's view, such an experience of multiple echoes misinforms the listener about the time and location of the explosion, but at the same time, it also informs her the approximate locations of the buildings. However, the listener should be able to become aware of something more. When she is fully aware of the fact that the

---

[17] O'Callaghan does not provide any source for this claim. A quick search in more than ten books in psychoacoustics fails to find any supporting evidence. Thus, this second feature should be taken with caution. Fortunately, the point I am going to make does not depend on it, so I just mention and leave it here.

[18] More detailed discussion of O'Callaghan's view on echo experiences will be provided in §5.2.6 when we discuss his disturbance event theory.

multiple distinct but related auditory objects are products of the same explosion, their successive appearance in her auditory experience constitute her awareness of the propagation of the compression wave. If this is accepted, then compression waves should be included in the list of auditory objects. Wave theorists can then reject eliminativism and maintain that we hear sounds.

Young might dig his heels in and maintain that although compression waves are auditory objects, they are not sounds. While this position is logically possible, it is very implausible. Wave theory seems to be the default position if compression waves can be auditory objects. This explains why almost all arguments against wave theory attempt to show that compression waves cannot be auditory objects by appealing to mismatches between their properties and the phenomenal character of auditory experiences. Without further argument from Young, eliminativism is no longer an option.

4.2.2 Subjectivism

Subjectivism flies in the face of the claim accepted in §3.2 that sounds are public objects. If subjectivism is true, the commonsensical idea that people in the same environment can hear the same sounds has to be given up, as sensations are available to their subjects only. While admitting that common sense cannot be any conclusive proof for or against a philosophical theory, subjectivists should provide very good arguments to defend their view. Unfortunately, Maclachlan (1989), the only contemporary subjectivist, fails to provide any persuasive argument.

Maclachlan’s argument can be analysed into four steps. First, he defends the austere view of the auditory world which says that we hear only sounds. Sound sources are inferred from the sounds we hear (ibid., pp. 8-9). So are reflective surfaces (ibid., pp. 20-22). Second, he rejects wave theory (ibid., pp. 26-28). Third, he further argues that sounds are sensations (ibid., pp. 28-30). Fourth, he explains why sounds appear to be public instead of private (ibid., pp. 30-31).

As we saw at the beginning of §2.2, the austere view either faces the identification problem of auditory objects or begs the question against some theories of sound. Therefore, even if it is true, we cannot employ it to support subjectivism. Moreover, Maclachlan does not argue for this idea but merely presents two examples. We can simply discuss one of them, as they are structurally the same. Suppose a burglar is moving about downstairs, and I claim that I know that because I hear him. Maclachlan suggests that I might have reason to believe instead that it is just a cat. A rhetorical question is then asked, "Isn't what we actually hear no more than … the noise made by the burglar moving about downstairs?" A rhetorical question, however, is no substitute for a proper argument.

Apparently, Maclachlan is trying to use his examples as intuition pumps, but I suspect he cannot get any unanimous support from people's intuition. In any case, we may expose the problem in his move here by identifying his implicit assumptions. At issue is my claim that I know that a burglar is moving about because I hear him. Maclachlan is right that this claim is fallible, as a cat might also induce an indistinguishable experience. Another way in which it can go wrong is that the *event* is misidentified. However, since the current issue concerns what *object* is heard, we may set aside this kind of error for the moment. The implicit contrast here is with my belief about the noise I hear. As I will show when I examine the second step of Maclachlan's argument, he thinks that such a belief is infallible. What makes the fallibility of a perceptual report to be relevant to the question concerning what is heard is the assumption that I cannot be mistaken about what I hear. However, this assumption is insufficient to establish Maclachlan's sonicist conclusion.

Consider this claim: I know that the source of the noise I hear is moving about because I hear it. This claim is infallible regarding which object is heard, as "the source of the noise I hear" picks out whatever it is that produces the noise, and all sounds have sources. Maclachlan's assumption therefore allows that I can hear the source. Therefore, he fails to show that I cannot hear anything other than the noise based on this assumption.

We granted that I might be mistaken when I claim I hear the burglar moving about, and the reason is that a cat can produce an indistinguishable experience. Maclachlan seems to be assuming that the indistinguishability of the experiences produced by a burglar moving about and a cat moving about implies that they represent the same object. Since neither the burglar nor the cat is represented by *both* experiences, so they cannot be the represented object.

The problem with this assumption is obvious. As I repeated several times in §2.2.2.3, indistinguishability is not identity. A burglar moving about and a cat moving about can produce two indistinguishable but non-identical auditory experiences, and there is no reason to suppose that the two experiences represent the same object. The kind of mistakes to which Maclachlan appeals is more properly understood as a misidentification of what exactly the object of an auditory experience is. This understanding requires that I do hear an object because I can misidentify an object in an auditory experience only if I hear it.

Since Maclachlan bases his claims that sound sources and reflective surfaces are inferred on the austere view of the auditory world above, and they are now shown to be ungrounded, I will not discuss them further.

Maclachlan argues against wave theory in his second step. He uses the Doppler effect as his main example. It is assumed that if wave theory is true, then the pitch of the sound heard should match the frequency of the compression wave. In the case of the source motion Doppler effect, i.e. the Doppler effect induced by a moving sound source when the observer is stationary, there is an objective change to the frequency of the compression wave, and the observer hears a pitch change as well. However, in the case of the observer motion Doppler effect, i.e. the Doppler effect induced by a moving observer when the sound source is stationary, the compression wave does not change, but the observer hears a pitch change. Maclachlan concludes from this second case that the sound heard is not the compression wave because the pitch of the former does not match the frequency of the latter. Rather, the sound heard is a joint product of the compression wave and the movement of the observer, and hence is an effect of the compression wave.

This argument begs the question. The perceived pitch is a quality of the sensation. Only if we have assumed the intended conclusion that sounds are sensations could we then take the experienced pitch as a property belonging to the sound heard. This question-begging assumption can be seen in other examples mentioned by Maclachlan, especially the case of "ringing in the ears", where he explicitly says the "the ringing *sound* we hear is merely an auditory sensation" (ibid., p. 28, my emphasis). It should be better to say that the ringing is merely a hallucination, and a hallucinated sound, of course, has no matching compression wave. Similarly, the case of the observer motion Doppler effect should be an auditory illusion. Illusions are, roughly speaking, cases in which the experienced quality and the actual quality of an object do not match. If we accept Maclachlan's argument, then we would need to give up the possibility of any auditory illusion. This is why I said above that Maclachlan assumes that my belief about the sound I hear is infallible. Apart from the problem of question-begging, this consequence is absurd enough for us to reject his argument against wave theory.

In the third step, Maclachlan tries to show that both hearing and the sense of pain have the same basic structure. By analogy, since pains are sensations, sounds should also be sensations. Just like hearing has sounds as its direct objects and sound sources as its inferred objects, the sense of pain also has pains as its direct objects and the causes of the pains as its inferred objects. For example, he thinks that we can feel a mosquito in the sense that we can infer it from the bite we feel directly, just like we can hear a dog in the sense that we can infer it from the bark we hear directly. Since we have already rejected the austere view of the auditory world, we should also reject this argument by analogy.

The fourth step is an explanation of the alleged mistake in thinking that sounds are public entities. As Maclachlan does not provide any good argument for subjectivism, we just do not need such an explanation. It is not a mistake to think that sounds are public entities after all.

I have examined each step in Maclachlan's argument for subjectivism and showed that it fails to show that sounds are sensations. We should therefore be objectivists of sound.

4.2.3 Many Sounds Theory

From the methodological perspective, many sounds theory is not an attractive option, as it is the least parsimonious theory on offer. Why should we identify sounds with multiple categories of entities if one is already enough? Proponents of this theory therefore bear a very heavy burden of proof. They need to argue against alternative theories by showing that they cannot cover every relevant phenomenon, and that no single category of sounds can do so.

Kulvicki (2017) proposes the only example of many sounds theory, but he fails to meet this requirement. He does not offer any objection to alternative theories. Rather, he merely motivates his two new accounts of sounds based on what he calls "audible profiles" (ibid., p. 87). I will discuss audible profiles in §4.2.4.1. For the moment, we can focus directly on Kulvicki's many sounds theory.

Kulvicki adopts the quality-splitting approach which distinguishes two sets of qualities—intrinsic and perspectival—belonging to objects, events, and environs. It is in sharp contrast with the more conventional object-splitting approach which splits the objects of auditory experiences into ordinary objects, events, etc. on the one hand and sounds as some distinct entities on the other hand.

Kulvicki's many sounds theory can be considered as an extension of his earlier position, i.e. his dispositional object property theory proposed in his (2008a). The most significant difference between these two theories is the addition of perspectival qualities into the list of sounds. A proper appreciation of the perspectival aspect of hearing and the related complexity of our auditory world is long due. Kulvicki's view is therefore commendable in this regard. However, it is not clear why these facts about our auditory perception could motivate a many sounds theory.

Kulvicki seems to think that identifying sounds with any single category cannot properly reflect the complexity of our auditory world. Therefore, we should call more categories "sounds". This is, however, reducing the question "What are sounds?" into a question merely about how we use the word "sound".

> This is perhaps an infelicitous use of the term. It's just odd to suggest that much of what we hear is not sounds, but something else. Instead, it's better to recognize that the category of sounds is a mongrel and to focus our efforts at understanding the many kinds of things heard. (Kulvicki, 2017, p. 92)

Being odd can hardly be counted as a reason in philosophical debate. It is perfectly fine to recognize the multiple categories of auditory objects while not calling all of them "sounds". Kulvicki owes us a proper argument.

4.2.4 A-spatial Theory

We mentioned two versions of a-spatial theory: one by Kulvicki (2017) and one by Strawson (1959). This section criticises them in turn.

4.2.4.1 Kulvicki's Audible Profiles

As we have just seen, Kulvicki highlights the perspectival aspect of auditory perception in his (2017). He attempts to explain this perspectival aspect by proposing what he calls "audible profiles". Audible profiles, according to Kulvicki, are abstractions over intrinsic and relational features of ordinary objects, events, and environs (ibid., p. 90). Such abstractions are objective features of the world. Kulvicki further suggests that one possible account of sounds is to identify sounds with components of audible profiles (ibid.).

Questions arise for this proposal. To begin with, it does not fit with our accepted claim in §3.6 above that sounds are caused by sound sources. Abstractions are neither causes nor effects. Instead, it is their instantiations which can be caused. Kulvicki might reject our causal claim about sounds, or instead amend his proposal by identifying sounds with concrete instances of audible profiles.

The next issue is to explain the attribution of those perspectival features. Audible profiles are joint products of ordinary objects, events, and environs. Among the cases mentioned by Kulvicki, perspectival features are attributed to either one of these three items, rather than all three together as a collection. This selective attribution calls for an explanation.

One possible explanation appeals to the similarity between audible profiles which share the same element. For example, holding the object constant and varying the event and the environ results in audible profiles which are similar in a certain way. This procedure allows us to identify the contribution of each element to audible profiles, and we just attribute the perspectival feature to the most crucial element.

This explanation invites us to question the motivation for proposing the notion of audible profiles. If at the end perspectival features are attributed to individual elements rather than collections of these elements, then why not simply consider those elements individually from the very beginning? In a game of tennis, every movement of the ball is a joint product of the intrinsic features of the ball, the racket, the player, the player's strike, the environ (e.g. windy or not), but there is no need to involve any abstraction over these elements to explain anything in the game. All we need are those individual elements. Why would the case of auditory perception be different? A further explanation is needed.

A third issue concerns the possibility of illusions of perspectival features. Kulvicki proposes audible profiles as the auditory counterparts of visual figures. An example of a visual figure is the elliptical shape of a coin viewed from an oblique angle. Just like audible profiles, visual figures are also perspectival—they are jointly determined by the intrinsic features of objects and the points of view. One thing not mentioned by Kulvicki is how optical illusions like the bent appearance of a stick half-immersed in water are understood in terms of visual figures.

Kulvicki cites some authors who have proposed similar ideas, but they have different opinions regarding the stick in water. We can focus on two of them. Hyman (2006, p. 75) describes the case of the stick in terms of *occlusion shapes*.

Roughly, an occlusion shape of an object is the shape of the smallest area one has to mark on a pane perpendicular to one's line of sight to cover that object from a point of view. Hyman makes it very clear that occlusion shapes are affected by the refraction and reflection of light, so the half-immersed stick has a crooked occlusion shape (ibid., p. 77). Therefore, there is no optical illusion of occlusion shape.

In contrast, Hopkins (1998, p. 62) says that we misperceive the *outline shape* of the half-immersed stick. He (ibid., p. 55) defines "an object's outline shape at a point as the solid angle it subtends at that point". The stick should have the same outline shape with a straight stick not immersed in water, but we misperceive it as sharing the outline shape with a bent stick.

We may say that the crucial difference between Hyman's occlusion shapes and Hopkins's outline shapes is that occlusion shapes are defined in terms of a procedure determined (in part) by the physical properties of light, while outline shapes are determined purely geometrically. However, this characterisation of the difference leads to a problem: what kind of geometry is assumed by Hopkins? Is it Euclidean geometry? Or the non-Euclidean geometry of space-time? This is related to how relevant the transmission of light is to the determination of perspectival features. One consequence concerns how celestial bodies like constellations should be treated, as they might involve light rays travelling along the curvature of space-time, and hence Euclidean geometry would lead to the classification of the perception of such objects as illusory.

As for the case of sounds, it seems to be unreasonable to define perspectival features in terms of geometrical objects like occlusion shapes. While people debate about whether compression waves can be objects of auditory experiences, no one would deny the necessary causal role played by them. Is the propagation of compression waves relevant to the determination of perspectival features? Compression waves require a medium for their transmission. They diffract significantly, such that we can hear around corners. Defining perspectival features independently of acoustics would render too many of our auditory experiences illusory. However, if we allow the diffraction, refraction, reflection, superposition of compression waves to affect perspectival features, it is not clear why those perspectival features are not just the intrinsic features of the compression waves at the location of an observer. Perhaps all these presuppose the rejection of wave theory, but I will show in Chapter 6 the objections to wave theory can all be replied satisfactorily.

The three issues above are some general worries concerning Kulvicki's proposal. Now I turn to his three reasons to support it. First, Kulvicki (2017, p. 90) says that his proposal explains why sounds appear to be distinct from objects, events, and environs, while at the same time seem to be intimately tied to them. As perspectival features, sounds are properties (or property instances) instantiated by objects, events, and environs. Properties are distinct from but tied to their bearers. However, Kulvicki does not show that this is the best explanation.

Furthermore, do sounds seem to be distinct in the mentioned way? We rejected in §3.5 the claim that sounds appear to be distinct from ordinary objects. Our discussion there targets at the object-splitting picture of auditory perception, according to which sounds are understood as distinct objects of auditory experiences, and therefore does not apply directly to Kulvicki's quality-splitting proposal.

I accept that there is an apparent distinction between properties and property bearers in auditory experiences. I also accept that we hear both intrinsic and perspectival properties. Nonetheless, I suspect we cannot have any non-question-begging reason to privilege any one of these items as sounds. We have already seen the identification problem of auditory objects in §2.1.3. This means that the phenomenology of auditory experiences is at least compatible with the following two possibilities: sounds as property-bearers and sounds as perspectival properties of something else. Therefore, we should reject Kulvicki's claim that sounds seem to be distinct from objects, events, and environs.

Second, Kulvicki (ibid., p. 91) also claims that his account of sounds as perspectival properties can explain auditory surrogacy. Roughly, auditory surrogates are auditory objects which represent some other auditory objects. For example, a person, standing still, can imitate what she heard earlier when a sound source, say, a fire engine, moved away from her. She imitates the siren from loud to soft, but we know that the siren imitated does not change its loudness. In this case, the person is an auditory surrogate of the fire engine. A loudspeaker is, according to Kulvicki, a very effective surrogate in this sense (ibid., p. 90).

Kulvicki explains auditory surrogacy in terms of the similarity between the auditory profiles of the original and its surrogate (ibid., p. 89). People can identify what is imitated from hearing the surrogate because they can recognise such a similarity. In other words, auditory surrogacy is explained by the similarity between the perspectival features of the original and the surrogate.

This explanation should be rejected. The surrogate can be heard from different perspectives, so that it has different perspectival features relative to different perspectives. For the sake of simplicity, let us assume that the surrogate imitates perfectly, such that there is one perspective P_s relative to which its perspectival features are qualitatively identical to the perspectival features of the original relative to the perspective P_o from which it is imitated. However, we do not need to hear the surrogate only from P_s to recognise what is being imitated. In fact, in the case of auditory surrogacy, it is very rare if not practically impossible to find, within the audible range of the surrogate, a perspective from which the original cannot be recognised. In contrast, as the angle from which a visual surrogate such as a photograph is seen becomes more and more oblique, it is more and more difficult and eventually becomes impossible to recognise what is pictured.

The problem is then that, in many cases, for a perspectives P'_s which is different from P_s, there is a corresponding P'_o which is different from P_o, such that the perspectival features of the surrogate heard from P'_s is more similar to those of the original heard from P'_o than to those of the original heard from P_o. If we follow Kulvicki's explanation, it seems we should then conclude that we recognise the same object as a surrogate of a different auditory situation if it is heard from a different perspective. For instance, if the person in our previous example is running toward me instead of standing still, I would instead recognise a stationary original from her imitation. This is clearly not the case.

A more plausible explanation would say that, in this last example, I hear the intrinsic features of the surrogate, through which I recognise what is imitated. I hear the change in the intrinsic loudness of the person's imitation, though this change is offset by the change in her distance from me and thus the perspectival loudness does not change. Nonetheless, I can recognise that she is imitating a fire engine which was running away from her. The intrinsic loudness of the siren did not change, but the perspectival loudness got softer from the perspective of the imitator because of the increase in the distance between them. If similarity matters for auditory surrogacy, then it should be the similarity between the perspectival features of the original and the intrinsic features of the surrogate which explains our recognition.

That said, I do not mean that this explanation is the best and thus the correct one. I just want to show that a better explanation of auditory surrogacy does not support identifying sounds with perspectival features more than with intrinsic features. If sounds are perspectival features, intrinsic features are non-sounds. Accordingly, the above explanation would appeal to the similarity between a sound of the original and a non-sound of the surrogate. In contrast, if sounds are intrinsic features, perspectival features are non-sounds. The above explanation would then appeal to the similarity between a non-sound of the original and a sound of the surrogate. In both cases, the similarity is between a sound and a non-sound. The case of auditory surrogacy does not offer any reason for us to favour either case.

Third and last, Kulvicki (ibid., p. 91) accepts a phenomenological version of sonicism—there is a phenomenological sense in which sounds mediate our perception of non-sounds. If sounds are perspectival features, then this sonicist idea becomes that we hear objects, events, and environs by hearing their perspectival features.

I do not think that the general idea that we perceive property bearers by perceiving their properties is beyond any doubt. I also argued in Chapter 2 that sonicism should not be taken as a background assumption in the debate about the nature of sounds. However, I would not challenge Kulvicki's proposal with these points. Rather, I am going to show that coupling his account of sounds as perspectival features with sonicism would result in an implausible view of auditory perception.

Recall that according to Kulvicki, sounds as perspectival features are components of audible profiles, and they are abstractions over intrinsic and relational features of objects, events, and environs. This means that regarding the level of abstraction, perspectival features are more abstract than intrinsic and relation features. As a result, sonicism implies that less abstract features are heard by hearing more abstract features. It is, however, implausible that perception operates in this direction.

Notice that I do not deny that we can perceive abstract objects. Whether abstract objects can be perceived alongside concrete particulars is an issue in the philosophy of perception. My objection is neutral on this issue. Rather, my target is the mediating role played by our perception of more abstract entities in our perception of less abstract entities.

There are three possible relations between our perception of objects at different levels of abstraction. First, they may be perceived equally directly. Second, we perceive more abstract entities by perceiving less abstract ones. Third, we perceive less abstract entities by perceiving more abstract ones. I believe Kulvicki would accept the first one, as this is required for his second proposal which is our target in §4.2.3 above. The second one is the more conventional view. It might be false as a universal claim, but it can hardly be denied that it is true at least in some cases. It is plausible that perception involves certain processes of mental abstraction which allow us to arrive at more abstract entities from less abstract ones, though we might question whether such mental processes are phenomenally conscious. As for the third possibility, it is not clear what kind of mental processes would work in this direction. Proponents such as Kulvicki who accepts a phenomenological version of sonicism should say more to clarify how it is mentally and phenomenologically possible.

To conclude, Kulvicki's account of sounds as perspectival features is not supported by the phenomenology of auditory experiences. While I have no knock-down argument against it, I maintain that we have insufficient reason to take it as a plausible alternative.

4.2.4.2 Strawson's A-spatial Sounds

Strawson (1959, p. 65) maintains that sounds "have no intrinsic spatial characteristics". This claim is presented in the context of his attempt to show that a purely auditory world is a no-space world, and it is indeed unclear if Strawson has provided any argument for it. This can be seen from a reconstruction of Strawson's discussion.

Strawson frames his argument for the claim that a purely auditory world is a no-space world as an objection against another argument. His opponent suggests that if sounds possess spatial characteristics, then a purely auditory world is not a no-space world (ibid.). The fact that attributions of spatial characteristics can be done "on the strength of hearing alone" is taken as evidence for this suggestion. By "on the strength of hearing alone," Strawson means that the object of hearing is not perceived with the concurrent help of any other sensory modalities. We might reconstruct Strawson's target argument as follow:

> P1. If spatial characteristics can be attributed on the strength of hearing alone, then sounds have spatial characteristics.

P2. Spatial characteristics can be attributed on the strength of hearing alone.

C1. Therefore, sounds have spatial characteristics.

P3. If sounds have spatial characteristics, then a purely auditory world is not a no-space world.

C2. Therefore, a purely auditory world is not a no-space world.

Notice that Strawson assumes the austere view of the auditory world that sounds are the only objects of hearing, and it seems he also assumes that his opponent makes the same assumption (ibid.). We rejected this assumption back in §2.2, but we may grant it for the moment. This assumption is required by P1, otherwise spatial characteristics might be attributed to auditory objects other than sounds, such that the consequent would not follow from the antecedent.

Another assumption, also required by P1, is that the auditory experiences based on which such attributions are made are veridical in spatial terms. This assumption bridges the gap between the spatiality of auditory experiences and the spatiality of sounds as objects of auditory experiences. Strawson does not spell out these two assumptions explicitly, but I suppose he would agree that they are implicit in his target argument.

Strawson supports his claim that sounds have no intrinsic spatial characteristics by denying the auditory significance of spatial expressions such as "to the left of", "spatially above", "nearer", and "farther" (ibid.). Strawson makes two further claims to support this denial. First, he thinks that the fact that the spatial characteristics expressed by these expressions are attributed on the strength of hearing alone can be sufficiently explained by the correlations between auditory properties of sounds and our non-auditory experiences of those spatial characteristics (ibid., p. 66). Call this "the sufficient explanation claim". Second, he maintains that such cross-modal correlations are necessary for the attribution of spatial characteristics on the strength of hearing alone. Call this "the necessity claim".

Before assessing these two claims, let us consider an objection against the relevance of Strawson's claim that sounds have no *intrinsic* spatial characteristics to his target argument. As we can see above, the consequents of P1 and C1 do not focus on the *intrinsic* spatial characteristics of sounds, hence P1 and C1 are compatible with Strawson's claim if sounds have *non-intrinsic* spatial characteristics.

One way to block this objection is by explaining what it is for a spatial characteristic to be intrinsic. Strawson does not provide any explicit definition, but a plausible interpretation is suggested by the sufficient explanation claim and the necessity claim—intrinsic spatial characteristics are spatial characteristics attributable to sounds not in virtue of any correlation between auditory properties and our non-auditory experiences of spatial characteristics. Understood in this way, the objection above cannot defend Strawson's target argument, as P3 would become unacceptable if sounds have spatial characteristics only in virtue of cross-modal correlations, in which case nothing follows about a purely auditory world.

We shall then move on to assess the sufficient explanation claim and the necessity claim in turn. The sufficient explanation claim, indeed, does not show that sounds lack intrinsic spatial characteristics. The attributions of spatial characteristics to sounds need not have only one sufficient explanation. Strawson should, but does not, show that cross-modal correlations provide the best explanation for such attributions. At least, it seems the hypothesis that sounds possess intrinsic spatial characteristics which can be perceived on the strength of hearing alone is a simpler and to that extent better explanation.

This simpler explanation is incompatible with the necessity claim, as the latter denies the possession of intrinsic spatial characteristics, understood in the way explained above, by sounds. Recall that intrinsic spatial characteristics are attributed not in virtue of cross-modal correlations. So, if the necessity claim is true, i.e. cross-modal correlations are necessary for the attributions of spatial characteristics, then sounds cannot possess any intrinsic spatial characteristics. Therefore, the necessity claim can be viewed as supplementing the sufficient explanation claim by excluding the simpler alternative explanation suggested above. However, it seems Strawson does not argue for his denial of the possession of intrinsic spatial characteristics by sounds at the end, as the necessity claim begs the question by making the same denial.

Can we save Strawson's a-spatial view of sound from the charge of question-begging? The crucial point concerns how *intrinsic* spatial characteristics should be understood. It might be objected that the charge of question-begging can be avoided if we understand intrinsic spatial characteristics in another way, and hence a more charitable interpretation should adopt such an understanding. A suggestion of this sort can be found in Nudds (2001, pp. 210-215).

A caveat. Nudds moves quite freely from the spatiality of auditory experiences to the spatiality of sounds in his discussion. This is problematic, since localised auditory objects need not be sounds, and sounds—even if located in space—need not be localised in auditory experiences. Assuming auditory experiences are largely veridical can be helpful only if auditory experiences of sounds, rather than those of other possible auditory objects, are veridical in spatial terms. Interestingly, Nudds does notice this distinction between sounds and experiences of sounds in a note (ibid., p. 226, n. 7). It is unclear why he still makes such a transition so freely. We may reject Nudds's view solely based on this problem, but I will nonetheless grant it for the following discussion and expose other problems of his arguments.

Here is a very simplified sketch of Nudds's discussion. He starts by noting that Strawson draws a contrast between vision and hearing—visual experiences are, but auditory experiences are not, intrinsically spatial. Nudds then proposes his interpretation of what Strawson means in saying that auditory experiences are not intrinsically spatial. Granted the transition from the spatiality of auditory experiences to the spatiality of sounds, we may expect that Nudds would conclude that sounds have no intrinsic spatial properties based on the claim that auditory experiences are not intrinsically spatial. However, Nudds does not make this move, though we may well wonder what stops him from doing so. Anyway, his discussion proceeds by arguing that sounds have spatial properties only contingently, in the sense that they can have no spatial property at all. He thinks that this "fact" about sounds supports the earlier conclusion that auditory experiences are non-spatial.

It appears that we could focus on the last part on sounds, as our concern is with the nature of sounds. However, since we granted the transition from the spatiality of auditory experiences to the spatiality of sounds, I will follow the flow of Nudds's discussion and evaluate Nudds's argument from the non-spatiality of auditory experiences first.

Nudds's reason for denying that auditory experiences are intrinsically spatial is that we are not aware of empty space. This claim as presented in the paper we are discussing now is quite perplexing. On the one hand, he allows that we are auditorily aware of a region of space as a potential location of sounds (ibid., p. 212); on the other hand, he argues that we are not auditorily aware of empty places (ibid., p. 213). But is not an empty place a potential location of sounds? If so, it seems Nudds is saying that we are both aware of and not aware of such a place auditorily.

This puzzle, assuming that Nudds did not change his view, is clarified in his later work (2009). We are aware of a region of space between two sound sources, but this awareness tells us nothing about this region of space, not even whether it is occupied or empty (pp. 88-89). We are in this sense not aware of *empty* places.

Recall that Nudds is trying to explicate the sense in which visual experiences are, but auditory experiences are not, intrinsically spatial. Therefore, to draw this contrast, visual experiences should provide us with an awareness of empty places. Nudds adopts an example from Martin (1992, p. 199): we are visually aware of the hole in the middle of a polo mint as an empty place. Nudds (2001, pp. 213-214) clarifies that his claim is not that visual experiences and auditory experiences have different spatial structure, but that auditory experiences have a non-spatial structure: our lack of awareness of empty places implies that we cannot be aware of any relation between a sound and a region of space it could occupy. We are aware of a sound but not as occupying a region of space. In contrast, we are visually aware of an object as occupying a region of space.

I have two objections to Nudds's view. First, if we shift our attention from reflective objects to light sources, the alleged contrast between visual experiences and auditory experiences disappears. Looking at the starry sky at night, are we aware of the dark region as empty? Consider another example. Imagine a rod with a row of LEDs on it placed in a completely dark room. If only the two LEDs at the two ends are turned on, we can see two light spots. However, this visual experience is indeterminate regarding whether the region of space between the two LEDs are occupied or empty. Are we visually aware of empty places in these two cases? Whatever answer you give, by parity it should be the same as your answer to the counterpart question regarding our auditory awareness of empty places.

Again, we encounter an example of how problematic it is to ignore the important difference between reflective objects and light sources. This difference has its counterpart in the case of audition which constitutes my second objection: we can be auditorily aware of empty places via echolocation.

We do not only hear sound sources, but also reflective objects via the reflected compression waves. Human echolocation might be unfamiliar to most people, but it is a skill used by many blind people in their daily life to navigate their environment. Even for non-blind people, the presence of reflective surfaces has a significant impact on the phenomenal qualities of their auditory experiences. The difference between shouting toward a wall nearby and shouting on an open field is auditorily salient. Supa, Cotzin, and Dallenbach (1944) reports a classic study on human echolocation in which two blindfolded subjects with normal vision quickly acquired the ability to detect via echolocation a distant obstacle when walking toward it. They succeeded 25 times to stop walking before running into the wall within only 40 and 44 trials (ibid., p. 144).

It might be disputable whether echolocation experiences are phenomenally auditory, as reports of a feeling of facial pressure are often made and so human echolocation is also called “facial vision”. However, the behavioural evidence suggests that the subjects are aware of the role played by hearing, as shown by how they tried to improve their performance. For example, one subject in the study above shuffled her stocking feet vigorously over a soft carpet runner under one experimental setting (ibid., p. 148), and all subjects thrust their heads forward while walking as if they were straining to hear under another experimental setting in which their ears are blocked with multiple layers of materials (ibid., p. 170).

The ability of blind people to navigate the environment with echolocation shows that they are auditorily aware of the difference between occupied and empty places. Otherwise, they would not be able to locate and avoid obstacles on their way. This parallels Martin’s example of the hole in a polo mint. Putting the cases of light or sound sources and reflective objects together, it seems a better understanding of vision and hearing is that: (1) in the case of light or sound sources, our visual and auditory experiences do not represent a region of space determinately as occupied or empty apart from the locations of the light or sound sources; (2) in the case of reflective objects, occupied places are perceived by the reflected light or compression waves, while empty places are perceived in virtue of the awareness of the surrounding reflective surfaces.

Nudds’s argument for the contrast between visual experiences and auditory experiences is flawed because the difference between light or sound sources and reflective objects is ignored. If visual experiences are intrinsically spatial because they provide us with visual awareness of empty places in the cases of seeing reflective objects, auditory experiences should also be counted as intrinsically spatial to the same extent. Conversely, if auditory experiences have a non-spatial structure in the case of hearing sound sources, visual experiences should have the same non-spatial structure when we see light sources *only*. As a result, Nudds’s interpretation of Strawson's claim that auditory experiences are not intrinsically spatial should be rejected. Moreover, granted the transition from the spatiality of auditory experiences to the spatiality of sounds, Nudds's argument does not imply that sounds are a-spatial as maintained by Strawson.

We can now move on to Nudds's three reasons for saying that sounds can have no spatial properties. The transition from the spatiality of auditory experiences to the spatiality of sounds is more salient here, as Nudds repeatedly draws conclusions about sounds from how we experience them to be. Again, I grant this transition for the following.

First, Nudds claims that we can hear a sound without spatial property, such as a ringing in one's ears (2001, p. 214). He further emphasises that a sound without spatial property does not sound like a sound merely without determinate location. I do not have any knockdown argument against this phenomenological claim. The fact is that my experience is different: a ringing in my ear does not appear to have no spatial property. I can identify instantly which ear it is in, and it appears to be located somewhere inside my skull.

Relatedly, in a note (ibid., p. 226, n. 4), Nudds claims that we can imagine a sound without imagining where it comes from, but we can only visualise objects from some point of view. He thinks that this feature of our imagination also shows that sounds can appear as having no spatial properties. This claim about imagination is suspicious at best. A perspective might not be salient when we imagine a sound, but it is easy to imagine a change in the perspective from which the sound is imagined. We can imagine the sound moving closer or further away or coming from a changing direction. If it has no imagined location from the very beginning, it is impossible to imagine a change in perspective. Rather, what would happen is a change from no perspective to a varying perspective. This is, however, not what seems to be the case. That said, this is again a phenomenological dispute, and we should not put too much weight on it.

Second, Nudds says that we do not identify or individuate sounds spatially (ibid., p. 214). I cannot see how this claim supports the non-spatiality of sounds. One of Nudds's examples is that you can identify two dogs which are fighting outside in the street by non-spatial qualities of their growls and yaps. But what does this ability show? Even if sounds are spatial, it does not follow that we must identify or individuate them spatially. Compare the case of vision. Is it necessary for us to identify objects via their spatial properties? I have a blue pen and a red pen in my bag. In most situations, I identify them by their colours only. Following Nudds's reasoning, it seems visual objects are therefore non-spatial as well. This reason is implausible and cannot make any distinction between auditory and visual objects.

We may have misunderstood what Nudds says. Perhaps he is not pointing to how we actually individuate sounds in particular cases, but whether differences in spatial properties are neither necessary nor sufficient for us to identify or individuate sounds. Consider the possibility that differences in spatial properties are not necessary first. Nudds uses as an example the auditory stream effect. Bregman (1994, pp. 17-18) describes an experiment in which a sequence of six tones, labelled 1-6 according to their pitch from low to high, are presented repeatedly in the order 142536. As the speed of alteration gets faster, at some point the subjects hear two sequences: one consisting of the lower tones (1-2-3-) and one consisting of the higher tones (-4-5-6). Moreover, they cannot attend to both sequences at the same time. They can report the order of either three-note sequences, but not the actual order of the six-note sequence. The relevance of this effect here is that while the two three-note sequences share the same spatial properties, we nonetheless individuate them as two sounds. Therefore, differences in spatial properties are not necessary for us to identify or individuate sounds.

It seems the auditory stream effect can only show that the way we individuate sounds is a poor guide as to the identity condition of sounds. The original description is that a six-note sequence is heard as two three-note sequences when the rate of alternation is fast. However, the same case can also be described as two three-note sequences are heard as one six-note sequence when the rate of alternation is slow. It is very strange if not utterly implausible that the number of sequences should change with the rate of alternation.

Is audio playback by a single loudspeaker a better example? I can hear an orchestral performance in the playback and distinguish the sound of one instrument from those of other instruments. Nonetheless, all sounds come from the same loudspeaker. Also, since we have only one loudspeaker, there is no stereo effect, hence the sounds appear to be coming from the same location. So, we have a case in which differences in spatial properties is not necessary for the individuation of sounds.

This example, I believe, suffers the same problem. Although the compression wave coming from the loudspeaker is much more complex than the six tones in Bregman's experiment, there does not seem to be any non-question-begging reason to say that there are multiple sounds rather than one sound only. This example can be a good case study for how our auditory system parses the auditory stimuli into auditory objects, but whether those objects are sounds or something else is a question to which the phenomenon is neutral. We saw in §3.8 that describing the same auditory object in terms of "a changing sound" or "a changing stream of sounds" can be "alternative and equally acceptable methods of bookkeeping" (Cohen, 2010, p. 312). We may apply this lesson from diachronic bookkeeping to the case of synchronic bookkeeping—calling the same auditory object "a complex sound" and "a complex of sounds" may also be "alternative and equally acceptable methods of bookkeeping". At least, we should be open to both methods before we have established what sounds are on independent ground.

How about the possibility that differences in spatial properties are not sufficient for the identification or individuation of sounds? Nudds does not consider this case, but we can easily come up with a possible example. Consider the case of stereo loudspeakers. Two loudspeakers, one on my right and one on my left, are playing the same signal at the same time. Assume that their distances from my ears are the same. In this situation, my experience is that a single sound is heard as located at the mid-point between the two loudspeakers. Two sounds with different spatial locations are not distinguished into two individuals. Therefore, differences in spatial properties are not sufficient for the identification or individuation of sounds.

Again, this example does not show a difference between auditory objects and visual objects. Stereo systems achieve the designed effects by taking advantage of how our auditory system integrates auditory stimuli. To this extent, VR headsets induce depth perception in a similar way. A VR headset presents two slightly different pictures to each eye. Our visual system then integrates the two different retinal images to create a 3D image. Just like the case of stereo systems, our visual system does not represent the two images as two individuals despite their different spatial locations. Therefore, we should by parity also conclude that differences in spatial properties are not sufficient for the identification or individuation of visual objects. Visual objects and auditory objects are not different in this regard.

Third, Nudds contends that a world of sounds can be dissociated from the world of material objects, but a similar dissociation is not possible for objects of vision and touch (op. cit., p. 215). This dissociability claim, if true, can show the non-spatiality of sounds only if it is also assumed that if a kind of entities can be dissociated from the world of material objects, they are not spatial. This assumption is not obviously true. As we will see in a moment, there seems to be spatial entities which are not material objects. We may reject this assumption if such entities do not belong to the world of material objects.

In order to avoid begging the question, Nudds must not assume from the outset the intended conclusion that sounds are not spatial. One possible way for sounds to be spatial is for them to be material objects. Therefore, the dissociability claim is at least logically compatible with the possibility that sounds are one kind of spatial entities—may or may not be material objects—dissociable from some other kind of material objects. If, however, the dissociability claim just amounts to the idea that we can hear one kind of material objects while seeing and touching some other kind of material objects, this innocuous fact is far from sufficient to establish the non-spatiality of sounds. For instance, if a student sitting at the back of a large lecture hall hears the loudspeakers and looks at the lecturer, this does not follow that the loudspeaker has no spatial property. Substituting the loudspeakers with sounds in this example would not make the argument valid, unless the non-spatiality of sounds has already been established on some other ground. But then the dissociability claim is redundant. Therefore, Nudds needs to assume that sounds are not material objects. Without prejudging what exactly sounds are, I am willing to concede this much, as none of the existing theories of sound takes sounds to be material objects.

What does Nudds means by 'dissociation'? Nudds does not make it clear, and no concrete example is offered. Presumably, the objects of vision and touch which are supposed to be dissociable from sounds are sound sources. It seems reasonable enough to treat silent or inaudible objects as irrelevant to issues concerning sounds.[19] Therefore, when Nudds says that, in ordinary experiences, objects of vision and touch cannot be dissociated because they are 'one and the same' material objects (ibid.), I assume he is saying that sound sources, among other material objects, are objects which can be seen and felt. Accordingly, the dissociability

[19] At least for the current issue. The case of echoes, as well as related phenomena such as reverberations and echolocation, may probably require a different treatment.

claim amounts to the claim that sounds can be dissociated from sound sources. Obviously, the truth of this claim depends ultimately on the nature of sounds and their relation with sound sources, and it is dubious that this issue can be settled simply by considering our perceptual experiences. Nonetheless, I have granted the transition from experiences to objects of perception, so I will not take issue with that.

Sound sources are not the only objects we can see or feel. In order to establish the dissociability claim, Nudds needs to show that none of the other objects of vision and/or touch are sounds. This is, however, in conflict with some theories of sound. For instance, if sounds are the vibrations of sound sources (Casati & Dokic, 2009; Pasnau, 1999), then they can also be objects of vision and touch. Pluck a string, and its vibration—i.e. its sound—is there to be seen, felt, and heard. The sound has spatial properties too. You can locate it in the surroundings, and it has an extension matching that of the string. Moreover, our visual, tactile, and auditory experiences of the sound appear to form a unity. The vibration is a property or an event. It is closely related to but not itself a material object. If by ‘the world of material objects’ Nudds means the collection of material objects only, then sounds are not in such a world, but they nonetheless have spatial properties. If the world of material objects instead contains also entities closely related to material objects proper, then sounds are not dissociable from the world of material objects. In either case, sounds are spatial.

Consider another theory of sound. Wave theory identifies sounds with compression waves. There is a sense in which compression waves can be dissociated from sound sources—the former is distinct from and does not depend for their persistence on the latter. Compression waves also have spatial properties different from those of their sources. A compression wave spreads across a certain finite region of an elastic medium at any moment. Throughout its lifetime, the region it occupies normally expands until its energy runs out. In contrast, the sound source, assuming that it is not moving around, stays at the origin of the expanding region of wave propagation throughout the period. Although we cannot see compression waves, we can sometimes feel them. An example is the drums of a rock band. You can both hear and feel the beats. So, if wave theory is true, then sounds are both objects of hearing and touch. Although sounds can be dissociated from sound sources, it is doubtful whether we could dissociate the auditory and tactile aspects of sounds in any sense stronger than merely attending to our experience in a single modality. Imagine a rock concert in which you can only hear or feel the beats—it cannot be the same kind of experience rock fans enjoy. Compression waves are not ordinary material objects like tables and chairs, so it is open for dispute whether they are included in the world of material objects. If they are, then sounds as compression waves cannot be dissociated from the world of material objects. Even if otherwise, Nudds's conclusion still does not follow, as sounds are nonetheless spatial entities.

What the above discussion shows is that Nudds's third reason against the spatiality of sound is far from clear. If the dissociability claim means more than merely the selective attention to our auditory experiences, its truth would be evaluable only with some particular background theory of sound. However, in such a case, the spatiality of sound should be determined by that theory directly, rather than by the alleged dissociability from the world of material objects. We have seen two theories which render the dissociability claim false, and there are still others which I skipped for the sake of brevity.

Some theories of sound may support the dissociability claim. For instance, subjectivism may hold that sounds are auditory sensations which have no spatial properties and can be dissociated from the material objects we see and touch in the external world.[20] The trouble for Nudds is that the dissociability claim is not obviously true, so he cannot use it to ground such a theory of sound. On the other hand, the dissociability claim depends on some theories of sound to be true. So, in the absence of an independently established theory of sound, Nudds is not justified to conclude from the dissociability claim to the non-spatiality of sounds.

We have examined Strawson's argument and Nudds's attempt to defend the a-spatial account of sounds. Both authors fail to show that this is the best view. Although we do not have any knockdown argument against this Strawsonian a-

---

[20] A sensation can be located in the sense of being a certain mental state *in* the brain of the body of a subject. Being spatially related to material objects is sufficient for something to be spatial. However, I suppose this is not the intended sense of being spatial. After all, entities which are ordinarily counted as non-spatial, such as a thought or an idea, would also be counted as spatial if this sense is adopted.

spatial account, there is also no reason for us to accept this view, as there are better alternatives in the literature.

4.2.5 Proximal Theory

As I said in §4.1, proximal theory has no advocate in the literature. This section will simply describe two problems of this theory. Before that, I should point out that there is some grey area regarding where proximal stimuli are located. Depending on how you describe the transmission of pressure waves from eardrums to the hair cells in the cochleae, you might locate the proximal stimuli at different places involved in this mechanical process. However, the two problems of proximal theory do not depend on this indeterminacy, so let us assume that the proximal stimuli which are identified as sounds are vibrations of the eardrums.

The first problem is that proximal theory implies massive systematic errors of how many sounds are heard. Binaural hearing implies that in ordinary situations, there are always two proximal stimuli produced by a distal sound source. However, they are always merged into a single individual in our experience. Given that the implication of the existence of massive systematic errors is generally agreed to be a bad sign for a theory of sound, proximal theory should be rejected.

The second problem is that proximal stimuli cannot be public objects of auditory experiences. In one sense, the vibrations of the eardrums are public objects. Although it is practically impossible to hear the vibrations of other people's eardrums, we can imagine a tiny homunculus going inside someone's ear canal and hear the vibration of the eardrum. However, in this case, that vibration is a distal stimulus for the homunculus. Only the owner of the eardrum can hear the vibration *qua* proximal stimulus. In other words, the vibration *qua* proximal stimulus can only be a private object of hearing. Since we have accepted in §3.2 that sounds are public objects in auditory experiences, we should reject proximal theory because it conflicts with our accepted claim.

5 Theories of Sound (II): Distal Theory

Many philosophers of sound interpret our auditory experiences as representing a sound as distally located at or near its source. This phenomenological claim is probably the most common motivation for distal theory. If sounds are distally located, then our auditory experiences are veridical in this respect. However, many such philosophers accept as well that sound sources also appear in our auditory experiences. If I am right in §2.1.3 that what we take to be experiences of sounds may be experiences of sound sources instead, then distal theory is not supported by this phenomenological consideration. Such experiences would doubtlessly be veridical if sound sources are represented as distally located, because they really are.

Of course, my claim at §2.1.3 is merely that it is unclear whether the represented thing is a sound or its source. So, proponents of distal theory may well have interpreted our auditory experiences correctly. We have an indecision here, and this is enough for us to be sceptical about distal theory.

With this general doubt in mind, this chapter challenges two families of distal theory—object property theory in §5.1 and distal event theory in §5.2. As it will be shown, every distal theory faces serious problems, and hence shall be rejected.

5.1 Object Property Theory

This section challenges the two main versions of object property theory—occurrent object property theory and dispositional object property theory—in turn. After that, I will discuss a few general objections to object property theory.

5.1.1 Occurrent Object Property Theory

Pasnau (1999) motivates his occurrent object property theory by first showing a problem of what he calls “the standard view of sound”. The standard view consists of two claims. First, sounds are *the* objects of hearing. Second, sounds are waves, understood as properties of the medium. In other words, the standard view combines the austere view of the auditory world and wave theory, and therefore should be counted as a view both about hearing and sounds.

Pasnau argues that the standard view is incoherent: when we hear a sound, compression waves are everywhere in the air, but we locate the sound at its source (ibid., p. 311). At least one of the two claims of the standard view should be rejected, and Pasnau chooses to give up wave theory.

Pasnau's aim is to locate sounds at their sources. We have seen a few options to do so in the taxonomy presented in §4.1, so why does he choose occurrent object property theory over other alternatives? It seems his decision is motivated by an analogy between sounds and hearing on the one hand, and colours and vision on the other hand. He thinks this analogy is supported by the fact that both sight and hearing are locational modalities which inform us about their respective objects as well as the locations of those objects (ibid., p. 313).

Earlier in §3.3, I have already argued that this is a disanalogy. I will not repeat myself here. Since Pasnau bases his view on this problematic ground, the claim that sounds are properties is not sufficiently grounded.

Pasnau's view is problematic at a more specific level. Tying sounds to vibrations of sound sources excludes sounds involving non-vibratory sources, but there are many such sounds. Popping a balloon leads to a sudden release of pressure and thereby generates compression waves. However, the volume of air originally in the balloon does not vibrate.[21] Wielding a bullwhip makes a crack which is a small-scale sonic boom, but sonic booms are not produced by vibrations.[22]

---

[21] It might be objected that the balloon breaks down into pieces of rubber, and these little rubber sheets vibrate and produce a loud pop. This cannot be the case. Even if the rubber sheets vibrate, they cannot produce any loud sound. Those rubber sheets are small compared to the wavelengths of compression waves audible for us. When such a sheet vibrates, the compression on one side is quickly compensated by the rarefaction on the other side within a wavelength of the wave created, because the surrounding air can move freely around the sheet. As a result, the wave becomes very weak even at a few centimetres from the sheet. This cannot explain the loud pop we hear.

[22] Like the case of the balloon in Note 21 above, even if the bullwhip vibrates, no intense compression wave can be created to explain the loud crack we hear. The two cases can be explained in the same way. The fact that a vibrating bullwhip alone cannot produce any loud sound is evident if we consider the strings on an electric guitar. The strings are not attached to a soundboard as the electric guitar has none. When the electric guitar is not connected to an amplifier, no matter how hard you pluck a string, only the string vibrates, and the sound is tiny.

Thunders are generated when columns of air are ionised by flows of electricity. The air columns are heated up to 30,000 °C in a few microseconds (Heller, 2013, p. 142). The sudden expansion of the air columns creates shock waves which then become acoustic waves, and we hear claps or rumbles as a result. The air columns, nonetheless, do not vibrate. All these cases are soundless in Pasnau's view. Common sense might be wrong in these cases, but it is more reasonable to reject Pasnau's theory.

Perhaps sounds are some other properties, and occurrent object property theory need not be based on any argument from analogy. There may be some other ways to argue that sounds are properties. An example of such arguments can be found in Roberts (2017, p. 340).

Roberts begins with a Moorean fact: some watch is ticking. This is how the watch is being, and properties are ways of being. So ticking is a property of the watch. On the further assumption that if *S*-ing is a property then *S* is also a property, tick is also a property. Roberts then generalises the cases by appealing to a common presupposition that perceptible properties of the same kind belong to the same ontological category and concludes that sounds are properties.

A possible objection is that what it means to say that some watch is ticking is that some watch is making ticks. The ticks should not be properties of the watch, but rather some things produced by its activity.

Roberts (ibid., pp. 340-341) replies that this objection merely paraphrases the original statement of the Moorean fact. The Moorean fact—some watch is ticking—is true as stated. No one can deny it, and it is more plausible than any claim to the contrary. Its truth value should not be affected by paraphrasing.

I cannot see how this responds to the objection. Roberts's response copies how Schaffer (2009, p. 357) responds to a similar objection to a similar argument for the existence of numbers. Schaffer argues from the Moorean fact that there are prime numbers to the existence of numbers. The objection says that this Moorean fact is true only according to the fiction of numbers. The unqualified statement of the Moorean fact is false, although a proper paraphrase of it is true. Schaffer then responds by saying that the Moorean fact is true as stated.

There is a crucial difference between Roberts's response and Schaffer's one. Paraphrasing "There are prime numbers" into "There are prime numbers according to the fiction of numbers" is based on a philosophical theory of numbers. Call this a philosophical paraphrase. In contrast, paraphrasing "Some watch is ticking" into

"Some watch is making ticks" seems to depend on no more than the linguistic knowledge about how "tick" as a verb is defined by "tick" as a noun. Call this a linguistic paraphrase.

A philosophical paraphrase changes what fact is expressed by the same expression, while a linguistic paraphrase uses a different expression to express the same fact. This explains why Schaffer's response works, as the paraphrase he opposes changes the fact expressed by the statement "There are prime numbers" and thereby its truth value. However, this is not the case for Roberts's response. The objection does not assert that "Some watch is ticking" is false but "Some watch is making ticks" is true. They are both accepted as true because they are just the same fact expressed in different ways.

It is not obvious which expression better reflects the structure of the fact expressed. Even if we accept that the fact consists in the instantiation of a property by the watch, it is possible that the property instantiated is rather a relation to some other entity, namely, ticks. If ticking is such a dyadic property, it does not follow that tick is also a property. Roberts is, therefore, wrong to assume that if *S*-ing is a property then *S* is also a property, and his argument fails.

In the absence of a better argument, I conclude that occurrent object property theory should be rejected.

5.1.2 Dispositional Object Property Theory

Kulvicki (2008a, p. 2a; 2015, p. 206)[23] identifies sounds with dispositions of objects to vibrate in response to mechanical stimulation. The two most distinguishing features of this view are included in this simple statement. First, sounds are dispositions. Second, mechanical stimulation plays an important role.

Like Pasnau, Kulvicki also bases his dispositional object property theory on an analogy between sounds and colours. They both ignore the important distinction between light sources and reflective objects. Kulvicki uses the reflectance physicalist account of colour to guide his discussion (2008a, p. 2a). Accordingly, colours of objects are their dispositions to reflect light under illumination. Similarly, objects have sounds, which are their dispositions to vibrate when being thwacked.

[23] Kulvicki (2008a) uses the two-column layout. I put a letter "a" or "b" after the page number to indicate the left and right column respectively.

Moreover, light and compression waves are the means by which we see colours and hear sounds respectively (ibid., p. 3b).[24]

A problem quickly emerges concerning the two perceptual means. Light has colour. If sounds are tightly analogous with colours, then it seems that compression waves should have sounds. However, Kulvicki does not draw this implication, and it conflicts with his account, as he locates sounds at thing sources only but not in the medium. The analogy between sounds and colours is not as tight as Kulvicki wants.

Moreover, light has another role to play in Kulvicki's account. Reflective objects are stimulated by light. Kulvicki's description of an illuminated object as being "stimulated by light" (ibid., p. 4a) may seem odd for those who think that reflective surfaces merely passively reflect incident light. However, it fits naturally with the event view of colour which understands colours in terms of how electrons on reflective surfaces are excited by the energy of incident light (Pasnau, 2009). For the sake of simplicity, I follow Kulvicki in focusing on the more standard reflectance physicalist account in the discussion below.

The light reflected by an object is jointly determined by the incident light and the reflectance properties of the object. Reflectance properties are best revealed by full-spectrum light such as sunlight. By analogy, it seems Kulvicki would need to say that compression waves are stimulants to reveal sounds of objects. This idea may sound strange, but it is true in the case of resonance. Put two identical tuning forks next to each other. Hitting one of them produces a sine wave which stimulates the other one and thus causes it to vibrate as well. However, the second tuning fork is not stimulated by a full-spectrum compression wave, and the first tuning fork is not stimulated by a compression wave at all.

According to Kulvicki, the stimulant of the first tuning fork is a thwack (ibid., p. 4a). A thwack is a brief stimulus similar to sunlight in the sense that both can be analysed into frequency components which have a flat spectrum profile. In other words, they are both full-spectrum stimulants. In Kulvicki's view, however, making a good thwack requires two further conditions (ibid., p. 5b). First, thwack an object all over and at once. Second, thwack it not too hard and not too soft.

---

[24] More precisely, only visible light and audible compression waves can be perceptual means. With this subtlety acknowledged, I gloss over it for the rest of the discussion to simplify the discussion.

The first condition leads to absurdity. To understand the problem with thwacking an object all over, we need some basic knowledge of acoustics. An elastic object vibrates if an external force which deforms it is released. Its parts bounce back to their original positions but overshoot. They then bounce back from another direction and overshoot again. This back-and-forth movement of the parts constitutes the vibration of the object. The existence of damping forces such as friction slowly reduces the displacement. Eventually, the vibration stops, and the object returns to its equilibrium state. A crucial point here is that the vibration is set off by a *deformation* of the elastic object.

Most elastic objects have multiple modes of vibration. A one-dimensional object (e.g. a string) vibrates not only at its fundamental frequency (f_0) but also at frequencies which are integer multiples of f_0. Objects with more dimensions, such as a two-dimensional membrane or a three-dimensional block, have even more modes of vibration which bear more complex relations to the fundamental one.

Each vibrational mode has both nodes and antinode(s) of displacement.[25] For the sake of simplicity, consider a string. When a string vibrates, the displacement from the resting position of each point on the string is different. The actual displacement is the sum of the displacements of all vibrational modes at that point. For each mode, there are points which have no displacement. Such points are the displacement nodes of that mode. For example, the fundamental mode has two nodes, each at one end of the string, and the second mode has one more node at the midpoint of the string. The point with maximal displacement between two nodes is a displacement antinode. There is one antinode for the fundamental mode, and two antinodes for the second mode.

Plucking a string at different points excites it in different ways. A vibrational mode is not excited if you pluck the string at the displacement nodes of that mode. The closer the point plucked and the nearest displacement antinode of a mode are, the stronger is the excitation of that mode. Since the two fixed ends of a string are the only points at which every mode has a node, and there are no two points on each

---

[25] Displacement nodes and antinodes should not be confused with pressure nodes and antinodes. Pressure builds up when particles cannot move freely, therefore pressure maximises at places where there is zero displacement. In other words, a displacement node is a pressure antinode; a pressure node is a displacement antinode.

half of a string which have the same level of displacement in the fundamental mode, there are no two points on each half of a string which, when plucked, lead to exactly the same excitation of the string. One can selectively excite only certain modes by plucking the string at the displacement nodes of the unwanted modes. This is why timpanists strike the timpani head at a few inches from the rim, as this allows them to selectively excite only a set of harmonically related modes, such that pitched notes can be produced.

It is this fact that thwacking an object at different places excites it in different ways which motivates Kulvicki's advice to thwack an object all over, in the hope that every possible excitation of the object would not be missed (ibid., p. 5a). However, an all-over-thwack can only lead to a worse result.

Assume that the two ends of a string are fixed, such that an all-over-thwack would not simply move the whole string to a new location without deforming it. Every point between the two ends are then displaced, so every vibrational mode other than the fundamental one has at least one node be displaced. Recall that if a node is displaced, the corresponding mode is not excited. Therefore, an all-over-thwack can only excite the fundamental mode. Every object would give off a pure tone if it is thwacked all over, and hence they all sound alike. This is hardly what we would expect from a "good thwack".

Can we rescue Kulvicki's view by dropping this "all-over" requirement for a good thwack? No. A thwack at a different place excites an object differently. There is, however, no normative standard which can distinguish a good thwack from a bad one in this regard. Not even in the case of music, where the quality of sound matters much more than in ordinary life. Different timbres are desired in different contexts. A darker tone is good for a sorrowful passage but not a romantic one. However, it makes no sense to say that a darker tone is good *simpliciter*.

Furthermore, some percussion instruments are designed to play higher or lower notes by changing the position at which they are struck. The *bianzhong* ("編鐘", literally meaning "set of bells") is a kind of ancient Chinese percussion instruments. Two different notes, separated by a major third or a minor third, can be played on each bell depending on where it is struck.

How many sounds does each bell have according to Kulvicki's view? Does it have two dispositions to vibrate or one disposition which can be manifested in two ways? In other words, does it have two sounds or one sound which has two different manifestations? It seems the second option is better, as the two notes are produced by selective excitations of different subsets of the vibrational modes of a bell. There is thus no fundamental difference between this case and striking a table at different points. Since Kulvicki would say that a table has one sound, he should, therefore, say that each bell has only one sound, and the two different notes are just different manifestations of the same sound.

Such a characterisation of the *bianzhong* shows that we are more interested in the manifestation of an object's disposition than Kulvicki's account seems to suggest. According to Kulvicki's view, a musician's interest in the sound of her instrument is an interest in its disposition. However, when the musician plays the *bianzhong,* she is obviously interested in the *two* notes she can play on the bells. When she practices the instrument, she wants to consistently produce each of the two notes of each bell. It is the two notes which act as the individuals around which her actions are centred. She cares about the two manifestations more than the single disposition revealed. I believe this example shows clearly what is really at the core of our conception of sound, a core missed in Kulvicki's account.

Another problem of Kulvicki's account concerns the case of loudspeakers. Loudspeakers have sounds, but Kulvicki's account does not allow them to have the music played as their sounds. Kulvicki distinguishes between the sound an object *has* and the sound an object *makes*. If you want to hear the sound a loudspeaker *has*, you should thwack it. The thud produced is a manifestation of its sound (ibid., p. 6b). In contrast, when we play a piece of music through the loudspeaker, it *makes* sounds which it does not have. If the sound an object has is its disposition to vibrate at its natural frequency, then what is the sound it makes but does not have?

Kulvicki tries to show that it is common to find objects making sounds they do not have by considering the attack and decay patterns of objects in response to stimulants (ibid., p. 8a). When an object is thwacked, its vibration undergoes a transient state which is different from the steady state of vibration at the object's natural frequency. This transient state is determined by the material constitution of the object, so it would not change easily without modifying the object. We can then attribute to the object a disposition corresponding to such a transient response, and this disposition is also a sound it has.

Kulvicki does not make it very clear how attack patterns explain what the sounds an object makes but does not have are. It seems the idea is that when we listen to a piece of music from a loudspeaker, we are indeed hearing a series of transient responses of the loudspeaker to its stimulant. However, this explanation ignores the difference between free oscillation and forced oscillation.

When you thwack an object, the thwack deforms the object, and this sets the initial state. The initial state and the material constitution of the object then determine how it oscillates freely without further stimulation. In reality, the oscillation is damped, so it eventually stops. However, you can apply a periodic force to an oscillator to keep the oscillation going. It is this forced oscillation rather than the free oscillation which explains the case of loudspeakers.

Forced oscillation also consists of a transient state and a steady state. When a loudspeaker plays a pure tone, we are probably not aware of the transient state, as it normally lasts no longer than a few wave cycles. After that, the loudspeaker enters the steady state of oscillation, the frequency of which is determined entirely by the frequency of the driving periodic force. It is this steady state which figures prominently in our experience of sounds played through loudspeakers. Kulvicki's explanation is mistaken because it would instead say that we hear transient states all along.

Kulvicki gets it right that objects have dispositions as stable properties which determine how they appear in our auditory experiences. He is also correct that we can identify such stable properties through hearing. However, such facts can support his theory only if it is assumed that such stable properties are all we care about in auditory experiences and that we hear nothing but sounds. Otherwise, there is no reason to identify sounds with such dispositions. We may instead be able to hear objects and their properties along with other things. At the very least, we also hear the manifestations of objects' dispositions, and it appears that, as suggested by the example of the *bianzhong*, such manifestations are also plausible candidates of sounds.

To conclude, we failed to find any persuasive argument for dispositional object property theory. On the other hand, we saw that this theory is not congruent with acoustics, and the *bianzhong* offers a strong case against it. We should therefore reject this theory.

5.1.3 General Objections to Object Property Theory

This subsection discusses three objections to object property theory in general. The first one can be responded easily and hence is not really a problem for the theory. In contrast, the second and third objections are stronger, and object property theory seems to have not enough resources to reply satisfactorily.

First, sounds survive qualitative changes through time in a way that properties cannot (O'Callaghan, 2009a, p. 581). Squeeze a ball, the ball takes up a new property of shape, but its sphericality is gone. Similarly, while a note can slide from C5 to A4, no property of it survives the pitch change.

We have examined this idea of survivalism in §3.8. It is equally acceptable to say that what survives the pitch change is not a single sound but a stream of sounds (Cohen, 2010, p. 314). We may also distinguish discrete sounds and composite sounds, such that the latter but not the former can survive qualitative changes (Roberts, 2017, p. 343). These responses may not fit very well with the common way of describing our auditory experiences, but it is not surprising that common language usage fails to capture the metaphysical nature of sounds. Object property theorists have a satisfactory reply to this objection.

Second, properties cannot be caused by sound sources. We accepted in §3.6 that sounds are caused by sound sources, so if object property theory does imply that sounds cannot be caused, then we should reject it.

A straightforward reply is that, while properties as universals cannot be produced, objects can be caused to instantiate sounds as properties. Our concept of sound source would then need to be revised. When someone strikes a drum with a mallet, the striking causes the drum to instantiate a sound. The striking is the event source in the sense that it causes the instantiation of the sound. Normally we would say that the drum is the thing source. However, if we accept object property theory, saying so would mean that the drum causes itself to instantiate the sound, and this does not sound right. It would be better to say instead that the mallet causes the drum to instantiate the sound and hence to count the mallet as the thing source.

So far so good, but a problem emerges if we consider another case. If two cars run into each other, the collision is caused by both cars. One way to describe the case is to say that each one of the two cars is the thing source of the instantiation of a sound on the other car, but this implies that there are two sounds, one for each car. Therefore, people are wrong to say that the collision makes *a* loud bump.

Another way is to say that there is one sound instantiated by one thing source, which is the two cars taken collectively. The problem is then that there is no counterpart to the mallet in the previous example. We have to say that the two cars cause themselves to instantiate a bump.

In sum, object property theorists may need to either provide an account of causation which explains how an object can cause itself to instantiate a property, or accept that people fail to count the correct number of sounds in cases like a car accident. Such a consequence may not be fatal to object property theory, but it at least shows that its proponents have more work to do.

Third, it is unclear how object property theory can accommodate the perspectival aspect of auditory experiences. Kulvicki (2017, p. 92) criticises the two object property theories proposed in his (2008a, 2015) and Pasnau (1999) as being “deaf to perceptive”. Although I have rejected his two attempts to count perspectival features as sounds in §4.2.3 and §4.2.4.1, I fully admit the reality of this perspectival aspect of auditory experiences. It is worth to see what exactly Kulvicki means by saying that the two object property theories are deaf to perspective.

Unfortunately, I cannot find any hint of how Kulvicki’s earlier view accounts for perspectival features. Perhaps this theory is deaf to perspective in the sense that it says nothing about it.

This is, however, not the case for Pasnau’s view. Pasnau (ibid., p. 319) discusses loudness constancy and says that, when I move closer to a sound source, the sound does not become more intense, but the loudness as a subjective feature of my experience increases. We can, therefore, expect that Pasnau would say that the perspectival features of auditory experiences are subjective, and sounds as objective properties of sound sources are independent of those perspectival features.

As a result, if Kulvicki is right that Pasnau’s theory is deaf to perspective, he should have changed what he means by being “deaf”. A charitable interpretation may say that Pasnau’s theory is deaf to perspective in the sense that it denies the relevance of perspectival features to what sounds are objectively. In any case, it is unfair to put Kulvicki’s dispositional object property theory and Pasnau’s occurrent object property theory on a par regarding the perspectivality of auditory experiences.

The next question is whether Pasnau’s treatment of perspectival features is acceptable. I believe that it leads to a problematic consequence if we extrapolate it to the case of echoes. To begin with, notice that Pasnau denies that we hear compression waves. His objection to wave theory assumes that we hear sounds only and that sounds and compression waves have different spatial locations. It follows that we do not hear compression waves. Despite the high degree of correlation between properties of compression waves and perspectival features of our auditory experiences, that we do not hear compression waves implies that our auditory experiences do not represent them, and hence we cannot talk about their objective facts based on auditory experiences.

Let us then consider the case of echo experiences. A firework is launched into the sky and explodes. You hear the explosion and also an echo one second later. On the one hand, it seems we can say that there is a one-second delay between the arrival of the direct sound and that of the echo based on the experience. However, on the other hand, Pasnau's view does not allow talks about compression waves based solely on auditory experiences, so what we can say accordingly is just that there is a one-second delay between my experience of the direct sound and my experience of the echo. It is highly implausible that the content of such an echo experience is so impoverished that it does not allow us to say any objective thing about the delay between the arrivals of the direct sound and its echo.

We may now step back and ask if Pasnau's theory has done justice to the objectivity of the perspectival aspect of auditory experiences. Is it really the case that our talk about perspectival features merely report our subjective experiences? Are the changes in the intensity of compression waves resulting from the changes in the distance travelled never represented in our auditory experiences? I do not think so. More will be said about our experiences of compression waves in §6.6.

Maybe Pasnau should allow that we can hear compression waves in addition to sounds at least in some cases, but this removes a crucial element of his argument against wave theory. If his theory can only be defended against the above objection at the price of reviving its rival theory, then it is in trouble in any case.

This objection from perspectival hearing is not limited to Pasnau's theory. A major motivation for object property theory is the thought that auditory experiences correctly represent sounds as located at their sources. When coupled with sonicism, this thought leads to a tendency to downplay the role of compression waves in accounting for our auditory experiences. To this extent, this is also the case for other distal theories. Given the fact that perspectival features are partly determined by the behaviours of compression waves, such a tendency easily becomes an obstacle for object property theory to duly appreciate the objectivity of those perspectival features. I do not mean that this is an insurmountable problem for this theory, but its existing proponents have not provided a satisfactory account yet.

5.2 Distal Event Theory

The taxonomy in §4.1 lists five distal event theories: event property theory, vibratory event theory, disturbance event theory, event source theory, and secondary event theory. Except for the last one, the other four are closely related. This is different from the theories we have seen so far, as they are relatively isolated from each other. Therefore, a different strategy is adopted in this section. Instead of treating each of the first four theories individually, I try to connect them by showing how the problems of one theory are handled by another theory. Some general objections are then presented, and finally, I move on to discuss secondary event theory at the end of this chapter.

5.2.1 Vibration: Event or Property

Once the idea that sounds are located at their sources is taken seriously, it is tempting to tie sounds to the vibrations of thing sources. This is because such vibrations have properties such as frequency, intensity, and spectrum composition which have salient correlations with auditory properties of sounds. We have seen above Pasnau's occurrent object property theory which identifies sounds with such vibrations understood as *properties* of thing sources. Similarly, vibratory event theory also identifies sounds with such vibrations, but it instead takes such vibrations as *events* happening to thing sources. Two notes can be made regarding this difference in the ontological categorisation of vibrations.

First, properties are universals, while events are particulars. The idea that we can perceive abstract entities is mysterious if not false. It might be said that our perception of concrete instances of properties enables us to "perceive" those properties as universals. However, I doubt whether this is still perception. In any case, granting that we do perceive properties, occurrent object property theory puts sounds at a further remove from the subject than vibratory event theory does. If sounds are properties, we hear them by hearing their instances. In contrast, if sounds are vibratory events, we hear them, period.

Second, the two theories are motivated by different considerations and hence rely on very different arguments. Occurrent object property theory takes sounds as analogous to colours in the sense that both are sensory properties. As we have seen above, this theory relies heavily on an argument from analogy. Since the analogy between sounds and colours is a bad one, this makes occurrent object property theory highly questionable. In contrast, vibratory event theory takes the temporality of sounds as a crucial desideratum. Sounds start, develop, and end in time. This makes them event-like. Similarly, the vibrations of thing sources are event-like in the same sense. Together with the observed correlations between sounds and objects' vibrations, it is natural to take them as the same events. Since vibratory event theory does not rely on any argument from analogy, the disanalogy between sounds and colours is not a problem for it.

While vibratory event theory performs better than occurrent object property theory in not relying on a bad analogy, it cannot escape another problem of the latter theory. Recall the three examples of non-vibratory sounds: popping a balloon, wielding a bullwhip, and thunder. Vibratory event theory also has to deny that these are sounds, and hence it is very objectionable.

There is, however, one possible defence of vibratory event theory. Although Casati and Dokic identify sounds with vibrations of thing sources, the discussion in their (2009) proceeds in more general terms. In short, their crucial idea is that sounds are non-relational events which can happen to an object in the absence of a surrounding medium. So, we might include non-vibratory events in the list of sounds. For example, when a balloon is popped, the sudden release of compressed air is a non-relational event, and this event is a sound. Similarly, when a column of air is expanded by a flow of electricity, the expansion is a non-vibratory event, and this event is also a sound. In both cases, we can identify a body of air which undergoes an event, such that it makes sense to say that the event is a sound which happens to a volume of air as a sound source.

Unfortunately, the case of bullwhip cannot be handled by this revised account. The crack created is a small-scale sonic boom, but what exactly is a sonic boom? When an object moves in the air, the air in front of it is compressed and forms an acoustic wave travelling at the speed of sound. As the speed of the object increases, the distance between each wave decreases. When the object breaks the sound barrier, the distance between waves reduces to zero and the waves merge. The extreme pressure resulted then forms a shock wave which is a sonic boom.

Unlike the two previous cases, there is no distinct volume of air which can be identified as the sound source. We can identify a region of high pressure at the front of the moving object, but this region is continuous with the rest of the surrounding medium, and hence should not be separated from the larger whole. In other words, the compression caused by the object's movement is an event happening to the surrounding medium directly. As a result, the sonic boom is not a non-relational event which can happen in the absence of a surrounding medium, and hence cannot be a sound even in the more general non-relational account of sounds.

5.2.2 Involving the Surrounding Medium

The case of sonic boom introduces a problem for theories which try to identify sounds with events which can happen in the absence of a surrounding medium. An obvious solution to this problem is to treat sounds as involving the medium. One way of doing so is to accept disturbance event theory which claims that a sound is an event in which an object disturbs the surrounding medium. An obvious illustration of the difference between vibratory event theory and disturbance event theory is the case of a tuning fork vibrating in a vacuum. There is a sound according to vibratory event theory, but disturbance event theory judges that there is none because there is no medium to be disturbed.

Disturbance event theory can explain sonic booms by saying that a supersonic object disturbs its surrounding medium in virtue of its supersonic movement. The sonic boom we hear is a sound identical to this disturbance.

For this explanation to work, we need to revise O'Callaghan's original version of disturbance event theory, as he maintains that sounds are not any disturbance events but only the periodic ones (ibid., p. 70). We should loosen this restriction and count more disturbance events as sounds, although it might not be a good idea to allow all disturbance events to be counted as sound. For instance, when I wave my hand gently, I disturb the air around me, but only some very strong principled argument would make it plausible to say that this disturbance is a sound. Anyway, for the present case, if we accept disturbance event theory, we can happily allow that the disturbance produced by a supersonic object is a sound. Disturbance event theory is therefore preferable to vibratory event theory.

Sonic booms give us just one single case for the medium-dependence of sounds. Opponents might then reply by treating this case as an exception. Even the more extreme response of denying that sonic booms are sounds might not be a very high cost for insisting on the non-relational view of sound. We might, therefore, look for a stronger reason to believe that sounds depend for their existence on the presence of a surrounding medium. Let us consider whether sounds can exist in a vacuum again.

Those who accept the existence of sounds in a vacuum like to draw an analogy between sounds in a vacuum and colours in the dark. However, commenting on the latter case, Pasnau (2009, pp. 367-368) points out that our intuitions are "notoriously divided, and so what one should want out of a theory of color is simply a principled answer that accounts for those divided intuitions." Perhaps the case of sounds in a vacuum is just the same. Kulvicki (2008b, p. 1115) thus concurs that "no answer to the vacuum question could settle matters over the nature of sounds."

We might, however, question whether these two cases are really analogous. When people talk about colours in the dark, they do not mean colours of light. Rather, they want to know if *reflective objects* still have colours when they are not lit. In contrast, in the auditory case, we are interested in whether *sound sources* produce any sound in a vacuum. Objects which just reflect compression waves are not relevant to the question. It is, therefore, a mistake to take the two cases as analogous.

Can we make a better decision on the case of sounds in a vacuum? O'Callaghan believes that we do have a good reason to think that no sound can be in a vacuum. His argument consists of three steps.

First, assuming that audition does reveal true auditory qualities of sounds, he points out that a sound is heard as possessing different auditory qualities if we change the surrounding medium. The same event would sound differently in the air or underwater (O'Callaghan, 2007b, p. 52).

Second, there is no non-arbitrary ideal medium for hearing sounds which is analogous to full-spectrum illumination in the case of colour. Full-spectrum illumination has normative significance because only it can reveal complete detail about surface reflectance properties, and hence it makes sense to attribute illumination-independent colour properties to objects. In contrast, although the same event would have different auditory appearances in different mediums, in no sense can we say that some medium reveals complete detail about the sound. Underneath the varying appearance, there is no invariant basis. Therefore, auditory properties of sounds are medium-relative and hence cannot be medium-independent (ibid., p. 54).

Third, in a vacuum, the absence of a medium means that no auditory property can be attributed to any sound. Since sounds cannot exist without any auditory qualities, so sounds do not exist in a vacuum (ibid., pp. 54-55).

This argument does not favour disturbance event theory specifically. At least, wave theory can also accommodate the medium-dependence of sounds. Since we are now discussing distal event theory, I will leave this alternative aside.

O'Callaghan argues for the medium-dependence of sounds by making a case for the stronger claim that sounds are medium-relative in the second step. The problem is not that he appeals to the bad analogy between colours and sounds. Although he mentions this analogy, it is not essential to his argument. Rather, the problem is that the evidence presented is inconclusive. The crucial observation concerns the change in the auditory properties we hear a sound as having when the surrounding medium is changed. This can be explained alternatively as an effect of the change in the medium on our perception. The auditory properties do not change but are merely perceived differently in different mediums.

It seems O'Callaghan has assumed that if auditory properties are medium-independent, then there should be at least one medium which can completely reveal them in our perception. This is not clearly the case. Consider the not entirely parallel case of 3D shapes. The 3D shape of an object is never completely revealed in one perspective. We need to view the object from multiple perspectives by moving around it or turning it around. It is, however, absurd to conclude that 3D shapes are perspective-dependent. Similarly, it is possible that auditory properties are medium-independent but are never revealed in full by one medium. If we hold that sounds cannot be re-experienced, then this would imply that no sound can have its auditory properties be completely revealed, as there is not a second chance to perceive it in another medium.

The effect a medium can have on our perception of sounds should have a physical explanation. O'Callaghan does not provide any such explanation. One plausible explanation is that compression waves propagate in different ways in different mediums. So, the medium affects our perception of sounds in virtue of affecting the transmission of compression waves. However, it is incoherent with disturbance event theory to allow the transmission of compression waves to determine the auditory properties of sounds as disturbance events. Indeed, when O'Callaghan discusses wave phenomena such as constructive and destructive interference and the Doppler effect,[26] he treats the differences in the apparent auditory properties as perceptual distortions caused by how compression waves propagate in the medium, and hence do not involve any real differences in sounds (ibid., p. 108).

If the transmission of compression waves is the only cause for the differences in the apparent auditory properties in different mediums, then this is compatible with the alternative I suggested. That is, auditory properties of sounds are medium-independent, but the compression waves which carry information about them propagate in different ways in different mediums. There is no standard medium which completely reveals the true auditory properties of a sound. Instead, every medium "colours" our perception of sounds in their characteristic ways, and only through varying the medium could we completely grasp the true auditory properties. However, assuming that sounds cannot be re-experienced, there is no chance for us to perceive the same particular sounds in different mediums. The best we could do is to experience instances of the same type of sounds in different mediums. That said, since most of us are terrestrial animals surrounded by the same atmosphere throughout our lifetime, practically we have no chance to grasp the true auditory properties by varying the medium. This is unfortunate for our theoretical interest, but for practical purposes, it is good enough to hear sounds only as filtered by air.

If this alternative picture is possible, then O'Callaghan owes us a further argument to support his claim that the auditory properties are medium-relative. Otherwise, his argument for the medium-dependence of sounds fails. One straightforward way he can go is to provide a physical explanation as to how changing the medium affects sounds as disturbance events. I take this as his burden of proof and hence I would leave it here.

In any case, we still have the case of a sonic boom to support the medium-dependence of sounds. Although such a single case might not be sufficient to

---

[26] More precisely, my current point only applies to his account of the source-motion Doppler effects, *viz*. the Doppler effects induced by motions of sound sources. When a sound source moves relative to the surrounding medium, the wavelength of the compression wave generated lengthens or shortens depending on the direction of the motion. This is a change in how the compression wave propagates in the medium. In contrast, the observer-motion Doppler effects are caused merely by the motions of observers, and hence do not involve changes in wave propagation.

silence the opponents, we might accept it for the moment, and move on to see a problem of disturbance event theory related to the distinction between disturbance events and transmission events.

5.2.3 Disturbance Events and Transmission Events

From the perspective of disturbance event theory, the distinction between disturbance events and transmission events is well-motivated. It enables us to separate features of proximal stimuli into those contributed by sounds as disturbance events and those contributed by the behaviour of compression waves as carriers of information. Sounds and their properties are independent of the transmission of compression waves. "[T]he effects of changes to compression waves as they travel through the medium do not alter the audible properties of a sound itself" (O'Callaghan, 2007b, p. 87). Disturbance event theory thus has a clear standard for determining the true auditory properties of sounds, and hence it can provide a determinate criterion for veridical perception.

Disturbance events are events in which new compression waves are generated and introduced into a medium; while transmission events are events in which pre-existing compression waves propagate, such as travelling across interfaces between different mediums, travelling through barriers, reflections, refractions, diffractions, etc. (ibid., pp. 97-98). The crucial mark of disturbance events is thus the generation of new compression waves.

Unfortunately, there are cases which cannot be handled by this distinction. Consider the oboe. The oboe is structurally a wooden cone (the bore) with the double reed as its tip and the bell as its open end. There is only a tiny slit between the two pieces of cane for the air to flow through. When the player blows into the instrument, pressure in the player's mouth rises and thereby closes off the slit abruptly. This leads to a sudden drop in air pressure inside the bore, which becomes a rarefaction travelling toward the bell. A small fraction of the energy carried by the wave escapes the bell, while the rest is reflected because of the impedance mismatch between the air inside and outside the instrument. Since waves change sign when they reflect at open ends, the rarefaction becomes a compression and travels back to the reed. It then reduces the pressure difference between the player's mouth and the bore, and hence the slit is opened. A new flow of air enters the instrument and reinforces the compression. A new cycle begins.

Like many people, O'Callaghan (ibid., p. 99) treats the column of air in the bore as a vibrating mass which disturbs the surrounding air. This suggests that the sound as a disturbance event is located at the open end. We might then wonder if there is any disturbance event at the reed. After all, it is where a compression wave is generated. Therefore, it seems both the reed and the bell are plausible candidates for the location of the sound we can hear. How should we decide? There seem to be three possibilities: at the reed, at the bell, or both.

First. If the sound is located at the reed, then we should treat what happens to the air column inside the bore as a transmission event. The compression wave generated at the reed reflects at or escapes the bell. Inside the bore, the wave reflects again and again, and a standing wave is generated. At the bell, the surrounding medium is not disturbed by the escaping wave. Rather, a portion of a pre-existing wave travels across the interfaces between the air inside and outside the bore. No new energy is introduced by the vibrating air column, and hence no new compression wave is generated by it. Only the airflow at the reed disturbs a medium and generates new compression waves.

So far so good. However, we should bear in mind the fact that the frequency at which the reed vibrates is partly determined by the frequency at which the air column vibrates. In other words, the disturbance event at the reed is partly determined by the compression wave caused by it. Opening a tone hole changes the wavelength and hence the frequency of the compression wave inside the bore, and this changes the frequency of the reed's disturbance event. This conflicts with the claim that sounds and their auditory properties are independent of transmission events. At least in the case of the oboe and many other woodwinds, the disturbance event at the reed and the transmission event in the bore are mutually dependent.

A possible objection would say that this is not a problem specific to disturbance event theory. For instance, a purely acoustic description of how the oboe works would also admit that the vibration of the reed is partly determined by the reflected waves in the bore. One way to explain away the apparent mutual dependence is to distinguish two levels of description. At a larger time frame, the vibration of the reed and the standing wave in the bore appear to determine the frequency of each other. However, at a more fundamental level, we can describe each flexing of the reed. Each flexing in the steady-state of vibration is affected by a reflected wave caused by the previous flexings. At this level, there are no two mutually dependent events. Instead, the disturbance event which sets up a transmission event is not identical to the one affected by that transmission event.

Disturbance event theory will run into another problem if it replies in this way. The above reply appeals to disturbance events which can at most be considered as individual impulses. Acoustically, an impulse signal contains many more frequency components than the ones corresponding to the pitch we hear. Therefore, it is not the right kind of event to serve as the physical ground for the perceived pitch.

In short, for disturbance event theory, the disturbance events which can explain the auditory properties of sounds are not independent of transmission events, while those which are independent of transmission events cannot explain those properties. The distinction between disturbance events and transmission events as drawn by O'Callaghan is thus incompatible with his theory of sound.

Second. If we locate the disturbance event at the bell, then the air column is a part of the sound source. This is plausible to the extent that the reed and the air column behave as a coupled system and can be described accordingly.

The problem is that sounds can be heard inside the bore. For the current option, the air column in the bore is not a disturbed medium. Instead, it is a part of that composite object which disturbs the medium external to the instrument. Since a sound as a disturbance event is perceived in virtue of the compression wave travelling in a medium, that the air column in the bore is not a medium means that no sound of the oboe can be heard inside the bore.

However, this is not the case. Recently, a new kind of microphone called "Intramic" has been released. It differs from conventional microphones in that it is inserted in the bore of a clarinet or a saxophone to record the sounds of the instruments. If the sound of an instrument cannot be heard in its bore, an Intramic cannot record it. It is technically difficult to put an Intramic in the bore of an oboe because of the structure of the instrument, but we can expect that sounds can be heard also in the bore of an oboe. Alternatively, we may imagine a homunculus sitting inside the bore of an oboe. Can she hear any sound? If sounds as disturbance events are located at the bell, then the answer is no. But then what does the homunculus hear if it is not a sound?[27] No plausible answer is available.

Third. Perhaps there are two disturbance events, one at the reed and one at the bell. A homunculus in the bore hears the one at the reed in virtue of hearing the compression waves bouncing back and forth in the bore. The audience hear the one at the bell in virtue of hearing the compression waves filling the concert hall.

This option is inconsistent with the distinction between disturbance events and transmission events introduced at the beginning of this subsection. All the energy escaping from the bell comes from the player's flow of air. No new energy is introduced at the bell, and hence there is no new compression wave generated there. All the waves coming from the bell are pre-existing waves escaped from the interface between the air column in the bore and the air external to the instrument. If there is a disturbance event at the reed, none can be at the bell.

We can anticipate an objection from O'Callaghan. He thinks that resonance is a kind of sounding induced by pre-existing sounds (ibid., p. 99). A resonator "actively disturbs the medium and does not merely passively transmit existing compression waves." He may then treat the air column as a resonator which actively disturbs the surrounding medium at the bell, thus there are in total two disturbance events located at the reed and the bell respectively.

In response, we should analyse O'Callaghan's understanding of resonance more carefully. He distinguishes actively disturbing a medium and passively transmitting compression waves. This distinction is basically the one between disturbance events and transmission events but spelt out in more detail—disturbance events are active while transmission events are passive. We would then want to know why these two kinds of events are characterised in this way.

What makes disturbance events active? Earlier we said that disturbance events introduce new compression waves. In some sense, this seems to fit the case of resonance well. Place a vibrating tuning fork near a resonator, and then stop the tuning fork abruptly. The resonator keeps on sending off compression waves for a short while. No doubt such compression waves would be absent without the

[27] Sonicists face a further problem: if the homunculus does not hear any sound, whatever she hears is not heard by hearing a sound, and thus sonicism is false.

resonator, and hence it seems they are introduced by the resonator. So, perhaps disturbance events are active in the sense that they introduce new compression waves.

But are there really new compression waves introduced by the resonator? Compression waves are a form of energy transmission. If there are new compression waves, then it means more energy is transmitted, but where could this extra energy come from? In the example just considered, no matter whether a resonator is present or not, the only source of energy is the strike on the tuning fork. If there is no extra energy, then how could there be any extra compression waves?

What really happens is that the energy carried by the compression waves reaching the resonator is absorbed by it and re-radiated. However, not all the incoming energy is re-radiated immediately. A portion of it is stored and reinforced by the energy carried by the next compression wave reaching the resonator. As a result, energy builds up in the resonator. When the original sound source stops vibrating, there is still energy stored in the resonator, which keeps on radiating to the surrounding medium in the form of a compression wave. The question is: are the compression waves coming from the resonator really new?

To count both the reed and the bell as medium disturbers, we need to count the compression waves coming from the bell as new. In other words, the compression waves from the bell and those from the reed are not numerically identical. However, this conflicts with O'Callaghan's description of other wave phenomena. For him, a wave maintains its identity after reflection from a wall and transmission across the interface between two mediums. At the bell, a pre-existing compression wave splits into two parts: one passes through the interface, while the other one is reflected. Both are transmission events which O'Callaghan does not count as introductions of new compression waves in other cases. It might be surprising to realise that one compression wave can split into two parts (or even more after multiple encounters with interfaces). Nonetheless, to maintain the consistency of his view, O'Callaghan should hold that no new compression wave is introduced at the bell.

The above consideration leads us to the difficult issue concerning the identity condition of compression waves, a problem which goes well beyond the scope of the present discussion. I will return to this issue in §6.5. All we need to see now is that O'Callaghan's view is inconsistent if we choose the third option that there are two sounds located at the reed and the bell respectively.

We have now examined three possible accounts of the oboe example, but none of them works. It is unclear where the problem lies. The distinction between disturbance events and transmission events looks theoretically neutral, so we might be giving up too much if we reject it. Perhaps it just needs a better analysis. Proponents of disturbance event theory would then have a special burden to clarify the distinction. However, if we do not accept disturbance event theory, it seems we need not be concerned too much. Fortunately, we do have other alternative theories.

5.2.4 Event Sources or Disturbance Events

Disturbance event theory has still another problem—it is phenomenologically implausible. To see the reason, we need some background.

Both vibratory event theory and disturbance event theory identify sounds with events distinct from event sources. Striking a drum causes the drumhead to vibrate and disturb the surrounding medium. The sound produced, be it a vibration or a disturbance, is not identical to the striking as the event source.

Such a distinction between sounds and event sources leads to the following questions. If both sounds and event sources are represented in auditory experiences, are they represented as related? If so, what is the represented relation? Furthermore, what is it like to experience the relation?

One possible answer, as proposed by O'Callaghan (2011a, p. 376), is that sounds are represented as constituting parts of their event sources. Striking a drum is an event source which has a disturbance event as a part, and this disturbance event is a sound. Casati et al. (2013, p. 464), in proposing their event source theory, criticises this suggestion as phenomenologically implausible. When I hear someone strike a drum, I hear an event represented as a distal happening. This distal event, however, does not appear to be a part of a larger event which is *also* represented in the same experience. The phenomenology of our auditory experiences is simpler and more unified than what O'Callaghan suggests.

The unity and simplicity exemplified by the events represented in auditory experiences are better explained by event source theory. By directly identifying sounds with event sources, there is no need to posit multiple events represented in auditory experiences. I hear the striking on a drum, and this striking *is* the sound I hear. This theory has a more parsimonious ontology, as it commits to the existence of one event where other theories posit two. Moreover, it allows direct auditory experiences of event sources, which are understood as equivalent to direct experiences of sounds.

I raised the identification problem of auditory objects in §2.1.3. When a coach is passing by, there is a concrete particular represented in my auditory experience. I said it is not clear whether that particular is the event of a coach passing by or the sound of a coach passing by. Event source theory provides a simple solution to this problem—both options are correct, because they are equivalent.

There is an aspect of the original version of event source theory which needs amendment. Event sources either involve the surrounding medium or not. The original proposal by Casati et al. (2013, p. 464) is that event sources can happen in a vacuum. In other words, event sources do not involve the surrounding medium. As we saw in §5.2.1, however, this independence from the medium suffers from the problem concerning sonic booms.

A non-vibratory object can fly at 320 m/s at different altitudes. This speed is subsonic at sea level but supersonic at altitude beyond 20,000 ft. In other words, the medium-independent event of flying at 320 m/s produces a sonic boom only if it happens at higher altitude.

If event sources are medium-independent, event source theory would need to choose between two options. First, there is a sound no matter at which altitude the event of flying at 320 m/s happens, even if no sonic boom is produced. Or, second, there is no sound no matter at which altitude the event of flying at 320 m/s happens, even if a sonic boom is produced.

Neither option is attractive. I stipulated that the flying object is non-vibratory, so there cannot be any other medium-independent event which can serve as an event source. Thus, the first option asserts the existence of a sound which cannot be identified with any event source. Such a sound is, therefore, not covered by event source theory. As for the second option, it is blatantly implausible.

Notice that there is no third option which says that the medium-independent event is or is not an event source depending on whether a sonic boom is produced. It is because the sonic boom exists in the medium, such that event sources would be medium-dependent if they depend on the presence of sonic booms. This contradicts the assumption that they are medium-independent.

It might also be suggested that the event source is not the event of flying at 320 m/s, but rather the event of breaking the sound barrier. Again, this latter event is not medium-independent, as there is no sound barrier in a vacuum.

The easy way out is to revise the event source theory into requiring that event sources should be medium-involving and hence medium-dependent. Instead of the event of flying at 320 m/s, the event source and hence the sound in the sonic boom case should be, say, the event of flying at 320 m/s *in the air at 20,000 ft*. Similarly, striking a tuning fork is not an event source, but striking a tuning fork *in such-and-such a medium* is.

A new worry is that the updated event source theory may suffer from problems facing disturbance event theory. Maybe. Even so, if the distinction between disturbance events and transmission events can be clarified, then both disturbance event theory and event source theory would be rescued. Nonetheless, event source theory would still imply a more plausible phenomenology than the one offered by disturbance event theory, and hence it is still preferred.

An interesting consequence of event source theory is that a sound can be perceived via non-auditory modalities, such that it can be perceived by a deaf person. Leddington (2019, p. 623) objects that if event source theory allows that a deaf person can perceive a sound, it seems the theory has unwittingly changed the topic.

A quick response. A deaf person still cannot *hear* a sound. They just perceive it via some other perceptual modalities. Why should this be objectionable? The only reason would be the traditional idea that sounds are proper objects of hearing. However, I have already rejected it in §3.1. If event source theory is true, it does not change the topic but teaches us something new about how sounds can be perceived.

There are two more troublesome objections though. First, the claim that sounds are event sources is itself paradoxical. A sound as an event source is caused by some other event. Should this latter event be counted as a sound source because it is the source of the sound? If so, then it is another event source. However, event source theory implies that it is then also a sound. Again, this sound has another event as its cause, and so we can identify another event source which is another sound. It seems this can go back to the beginning of the universe, and there seem to be events which cannot be plausibly identified as sound. For example, if my seeing a written instruction causes me to make an announcement in front of the whole class, then my seeing the instruction is a sound. This is unacceptable.

Alternatively, we can describe the problem in the opposite direction. A sound is an event source of another sound, which is, in turn, an event source of still another sound. This goes on infinitely, such that any successful attempt to produce one sound results in an infinite number of sounds.

This problem cannot be solved by allowing self-causation, as a sound's coming into existence would not be explainable. A more promising solution would be to specify the events normally picked up by the term "event source" in some other way. One possibility is to distinguish two kinds of event sources: one for sounds and one for compression waves. Event sources of sounds produce event sources of compression waves, and these latter events are sounds.

This revised event source theory, however, appears to be just a disguised version of disturbance event theory. Arguably the most proximal causes of compression waves are disturbance events, thus identifying sounds with event sources of compression waves is equivalent to identifying sounds with disturbance events.

Second, the duration of an event source does not correspond to that which appears in our auditory experience—it is generally too short.[28] For example, plucking a string is an instantaneous event source, but the event represented in our auditory experience appears to be much longer. Indeed, the duration of this longer event corresponds to the duration of the string's vibration. Moreover, there is a neat correlation between how this represented event unfolds and how the vibration of the string changes throughout the period. The same correlation exists between that experienced event and the disturbance of the surrounding medium caused by the string.

Identifying sounds with event sources would then have either one of the following two possible consequences. First, it may mean that we hear sounds in addition to that longer event represented in our auditory experience, regardless of whether it is a vibratory event or disturbance event. This is in effect putting event source theory on a par with disturbance event theory in terms of the implied multiplicity of events represented in auditory experiences.

Earlier above, we noted that disturbance event theory implies that both a sound and its event source are represented, and this does not square with the unity

[28] Nudds (2015, p. 282) also points to this mismatch in criticising distal event theory.

and simplicity of the event represented in our auditory experiences. It was then claimed that event source theory is preferred because it can avoid this implication. Now, what we have just seen is that this is not true. Event source theory also implies that two events are represented in auditory experiences. The two theories therefore face the same phenomenological objection.

The second possible consequence of event source theory is that our auditory experiences of the durations of sounds are systematically illusory. The illusion may be explained in terms of how a thing source responds to the event stimulating them. The stimulating event is an event source, i.e. a sound. As for the response of the thing source, it is a separate event which normally lasts longer than the event source. In hearing the duration of a sound, the duration of the response event is mistaken as the duration of the event source, and hence the experience is illusory.

This time, event source theory can maintain the unity and simplicity of the events represented in auditory experiences, but only at the cost of embracing an error theory of auditory perception. Sounds do not appear as they really are, and hence auditory phenomenology does not support identifying them with event sources. The initial appearance of event source theory as being more faithful to the phenomenology of auditory experience is then counterbalanced, if not overthrown, by this implication. Again, we have no reason to prefer event source theory over disturbance event theory.

5.2.5 Sounds as Event Properties

The failure of event source theory shows that we should better keep the distinction between sounds and event sources. At the same time, we may still want to avoid unnecessarily multiplying the events represented in auditory experiences. Event property theory can be viewed as an attempt to achieve both aims.

By identifying sounds not directly with event sources but with their properties, event property theory maintains the distinction between sounds and event sources. By not identifying sounds with events, event property theory does not imply that we hear events other than event sources. Sounds are "fused" or "united" (Leddington, 2019, p. 627) with event sources by being their properties, and hence the apparent unity and simplicity of the events represented in auditory experiences are preserved in this theory.

Although event property theory is preferable to event source theory, it also faces the same objection that event sources are in general shorter than the events represented in auditory experiences. Plucking a string happens in an instant, but what we hear lasts for, say, three seconds. How long does the sound last? If it lasts for three seconds, then how could an instantaneous event possess a property outlasting it? If it lasts for just an instant, then what is that three-second thing represented in my auditory experience? Is it some other event we hear, and if so, why do we seem to hear one unified event only? Or do we simply misrepresent the duration of the event of plucking the string and its sound? These are basically the same questions which we found no answer for source event theory, and there does not seem to be any satisfactory answer available for event property theory as well.

The fact that both event source theory and event property theory face the same objection above may suggest that what leads to the problem is their common idea of tying sounds to event sources. However, the more fundamental mistake of these two theories appears to be that they reverse the prominence of disturbance events and event sources in our auditory experiences. Tying sounds to event sources implies that event sources figure more prominently in our auditory experiences than other distal events, such as vibratory events and disturbance events. However, the objection above clearly shows that it is not the case. It is the string's vibration or disturbance, rather than the instantaneous event of plucking the string, which captures our attention and determines the qualities of our auditory experience. Experiential evidence, therefore, works against any attempt to tie sounds to event sources.

5.2.6 A General Objection from Echo Experiences

Disturbance event theory seems to be the best distal event theory so far. It does not tie sounds to event sources. It allows non-vibratory sounds and accepts the medium-dependence of sounds. Nonetheless, there are two unsolved problems. First, the identification of sounds with disturbance events, when conjoined with the distinction between disturbance events and transmission events, fails to give a satisfactory account of the sounds of wind instruments. Second, the thought that event sources are also experienced auditorily leads to the question of whether sounds as disturbance events are experienced as related to event sources, and a positive answer to this question can hardly be squared with the unity and simplicity of the phenomenal character of our auditory experience.

These two problems may not be fatal. The first one may be solvable by a better analysis of the distinction between disturbance events and transmission events. The second one may be avoided by rejecting that event sources are auditorily experienced. While I see no reason to block these two responses, both are not attractive. The distinction between disturbance events and transmission events looks clear enough, hence any attempt to further clarify it would likely be *ad hoc*. As for the assumption that event sources are experienced, it is more plausible than its contrary.

Apart from these two problems, I would like to discuss a problem, not only for disturbance event theory, but generally for all distal event theories we have discussed thus far. It is the same objection from echo experiences we discussed in §5.1.3 concerning object property theory. Let us look at how proponents of distal event theory understand echoes and echo experiences.

Both event source theory and event property theory have not been discussed concerning echoes and echo experiences. Casati and Dokic (2009, p. 99) simply dismiss echo experiences as clear cases of spatial misrepresentation. The only in-depth discussion is available in O'Callaghan (2007b),[29] so I will focus on it.

O'Callaghan (ibid., pp. 124-125) holds that an echo is a primary sound re-encountered, and an echo experience is just a distorted experience of the primary sound with distortions of its location in time and space as well as its auditory properties. There is a further condition that a *distinctive* echo experience, in which the echo is experienced as causally related to the primary sound, requires that there is an interval of a certain length between the arrival times of the direct wave and the reflected wave (ibid., p. 127).

To evaluate O'Callaghan's view, let us consider the following case of a distinctive echo experience. Suppose there is an explosion. One part of the generated compression wave goes straight to my location, while another part of it goes to a wall behind me, reflects, and arrives at my location one second after the arrival of the direct wave. Both the direct wave and the reflected wave cause me to have an auditory experience. Suppose then I say to my friend, "I hear an echo one second after I hear an explosion." How would O'Callaghan interpret this sentence?

---

[29] The account presented in O'Callaghan (2007b) is basically the same as the one in O'Callaghan (2007a). To simplify things, I reference the former work only.

To begin with, O'Callaghan can treat what I say as reporting some subjective fact about my two experiences of the very same explosion—i.e. that there is a one-second delay between my first and second experiences of the same explosion. But this does not seem to exhaust the content of my experience.

I contend that the delay experienced is as objective as any time gap between two independent audible events. In an alternative scenario, two explosions under appropriate circumstances may produce in me an indistinguishable experience of a one-second delay. It would be implausible to allow objective experiential content in this alternative case only but not in the echo case.

O'Callaghan may try to explain this objectivity in the experienced delay by saying that it represents a one-second delay between two distal events—the explosion and the reflection of the wave. Unfortunately, it would be too unlikely for such an experience to be veridical. A delay of very different length between the explosion and the reflection can also result in the same experienced length of it if the explosion, the reflective surface, and the perceiver are in a suitable spatial arrangement. For such an experience to be veridical, the explosion and the reflective surface has to be equally far away from the perceiver, with the distance not measured in spatial terms but in the time needed for the compression wave to cover it.[30] This condition is, however, rarely met. When I experience a one-second delay, I can only hope that I am lucky enough to be in such a rare situation such that my experience veridically represents the delay between the two distal events.

Most if not all people who are not born deaf have in their lifetime at least some instances of distinctive echo experience. However, this explanation implies that only a very tiny subset of such experiences of a delay is veridical. Most people would have nothing objectively true in their relevant experiences. This is implausible.

There is still another alternative explanation. Instead of treating the experienced delay as representing a delay between two distal events, this alternative says that it represents a delay between two proximal events—the arrival at my ears of the direct compression wave and that of the reflected compression wave. If I am

[30] Since the speed of compression waves varies with many factors in the medium, measuring the distance in spatial terms would tremendously reduce the generality of the veridicality condition mentioned in the text.

right about the objectivity of the experienced delay, this alternative seems to be a much better explanation than the previous one. Assuming that my auditory system is working properly, the length of the experienced delay is guaranteed to match the objective delay between the arrival times of the primary wave and the reflected wave. Even if the explosion and the reflection of the compression wave are not separated by one second, my experience would not thereby have nothing objectively true.

However, this alternative is unavailable for O'Callaghan. The two proximal events of the arrivals of the direct and reflected waves are not experienced in his view, since he denies that compression waves are objects of auditory experiences (ibid., pp. 69, 163), and a perceiver cannot experience the arrival of compression waves without experiencing the compression waves. Therefore, he cannot explain the objectivity of the experienced delay in this way.

I do not deny that my experienced delay most likely misrepresents a time gap between two distal events. All I need to show is that my experience has some remaining objective content which can only be explained if we allow the proximal events of the arrivals of compression waves to be experienced. This would be enough to show that O'Callaghan's view, and arguably other distal theories of sound, cannot afford a satisfactory explanation of the relevant experience. To achieve this aim, we consider another example.

Suppose I connect an electronic metronome to a loudspeaker placed at a certain distance from a large concrete wall on the opposite side of an open field. The metronome is set to 60 beats per minute. Standing right next to the loudspeaker, the echoes reflected from the wall are not synchronised to the ticks directly from the loudspeaker. I then move closer to the wall, such that the echoes can reach me earlier, while the ticks from the loudspeaker take longer to hit my eardrums. The intervals between the ticks and the echoes then narrow down, and finally they are perfectly in sync.

In this example, I do not merely synchronise my subjective experiences of the ticks and the echoes. If I have carried a microphone during the whole process, it can also record the synchronisation process between two different things. However, neither the loudspeaker's ticking nor the reflection of compression waves at the wall are changed by my movement across the field. Therefore, the things which are synchronised are not these distal events. Rather, they are the proximal events of the arrivals of the direct waves and the reflected waves. If O'Callaghan is correct that such proximal events are not experienced, no satisfactory explanation of this scenario can be offered. Therefore, we should reject O'Callaghan's view and allow that proximal events of wave arrival can be experienced auditorily.

To be fair, this conclusion does not necessarily show that O'Callaghan's theory of sound is false. After all, I have only shown that his view of our echo *experiences* should be rejected. Perhaps sounds are nonetheless distal events, and we experience proximal events in addition to them. However, a crucial step in O'Callaghan's argument for his disturbance event theory is the denial that compression waves are experienced. If it is allowed that compression waves are experienced alongside distal events, there is no obvious reason to prefer distal theory over wave theory. All the alleged experiential evidence, such as the perceived distal locations of sounds, can be characterised instead by saying that those are not locations of sounds as compression waves but the locations of the distal events experienced. Furthermore, to the extent that distal theories of sound suffer from the problems discussed in the previous subsections, if wave theory can avoid those problems, as I am going to show in next chapter, we should accept wave theory instead. Before that, I end this section with a criticism of a very different distal theory—the secondary event theory.

5.2.7 A Different Angle: Secondary Event Theory

The most significant aspect in which secondary event theory differs from other distal event theories is that it is anti-physicalist. This can be seen from both of the two main theses—that sounds are pure events and that sounds are *audibilia*. I shall show that both theses are implausible.

First. The claim that sounds are pure events is metaphysically unattractive. The existence of events which just happen but without happening to any object is highly suspicious, and hence require independent arguments. However, Scruton offers no such argument in his work. Instead, he just appeals to what he calls "acousmatic experiences" in supporting the claim that sounds are such events.

Acousmatic experiences of sounds are experiences in which sounds are experienced alone (Scruton, 1997, p. 11). Scruton does not deny that sound sources could in some sense be heard, but he holds that in experiencing sounds acousmatically, we miss nothing essential about the sounds (Scruton, 2009, p. 58). This suggests that sounds really can be completely detached from their sources, and hence they are events which happen not to any objects—i.e. they are pure events.

In §2.1.3, I introduced the identification problem of auditory objects. In having an auditory experience, it is not obvious whether the represented object is a sound source or its sound. It should be shown why the represented object in what Scruton refers to as an acousmatic experience is a sound alone rather than the sound source.

I expect Scruton would respond by appealing to the case of music. Probably the most well-known thesis in his philosophy of music is that genuine musical experiences are acousmatic. In listening to a live performance, we hear musical structure, and this structure can only belong to the sound but not to the instruments and musicians on the stage. Therefore, when we listen to music, the auditory object is the sound.

Such a response is far from persuasive. I have shown in §2.2.2.1 that event sources can bear auditory properties. Given the plausible assumption that musical structures are structures between auditory properties, Scruton's possible response above fails to show that the represented object is a sound rather than an event source. If it is an event source, it is an event happening to thing source(s). Therefore, the alleged "acousmatic" experience does not show that sounds are pure events.

Indeed, even if we grant that sounds can be detached from their sources, there is still no reason to accept Scruton's view. Wave theory, for example, also accepts the claim that sounds can be detached from their sources, but they are not pure events. Rather, wave theory, as I will argue in Chapter 6, can hold that sounds as compression waves are events happening to the medium.

Scruton would disagree. He does not merely reject physicalist distal theories. He also rejects wave theory. He thinks that identifying sounds with compression waves would eliminate the "phenomenal reality" of sounds (Scruton, 1997, p. 6). Just like red light—a light wave of a certain wavelength—might remain what it is even if it no longer appears red, compression waves remain what they are even if no one can hear any sound. Therefore, the wavelength of the light wave cannot explain the redness of light, and the presence of compression waves cannot explain what it is for there to be sound.

It is hard to see what exactly Scruton means, unless we assume the second main thesis of his view—sounds are *audibilia*. Consider redness first. Scruton assumes that colours are secondary properties. He thinks that we cannot capture what it is for light to be red in terms of the wavelength of light, because light with the specified wavelength can remain what it is even if it no longer appears red. Similarly, it seems equally possible that red light may have entirely different wavelengths. Therefore, the redness of light cannot be explained in terms of its wavelength.

By analogy, there can be compression waves even if no one can hear any sound. Also, despite its being physically impossible for there to be sounds without compression waves, this is not a metaphysical impossibility. The "ultimate fact" of there being a sound is just the counterfactual that if a normal observer is there, she would *hear* a sound in normal conditions (ibid., p. 7). The essence of sounds is therefore tied to their *auditory* appearance to normal observers in normal conditions. They are therefore *audibilia*—entities the essence of which resides in "the way they sound" (Scruton, 2009, p. 57).

Scruton then proceeds to show that sounds are not properties, therefore the argument above does not show that sounds are secondary properties of compression waves. I have nothing to say about this move. My objection will focus on his argument for the claim that sounds are *audibilia*.

The secondary property theory of colour has the problem of infinite regress (Levin, 2000, p. 163). An object is red if and only if it appears red to a normal observer under normal conditions. For our current purpose, we can put it in short—an object is red if and only if it is disposed to look red. If we apply this analysis to the second occurrence of "red", then it follows that an object is red if and only if it is disposed to look to be disposed to look red. Once again, the analysis can be applied to the second occurrence of "red". This process can go on infinitely, and thus we have an infinite regress.

Scruton (2009, pp. 53-54) tries to stop this regress by saying that we know what it means for an object to appear red through our own subjective experiences with red objects. As for the case of other people, we can attribute the same kind of experiences to them if they can discriminate red things by just looking at them.

Granted this response, I do not think that we can say the same thing in the case of sound. By appealing to the phenomenal character of experiences of red objects, Scruton has implicitly made the physical objects indispensable. Perceptual experiences are joint products of external stimuli and the operations of our perceptual systems. I can have a veridical visual experience of a red object only if there is an object which stimulates my visual system. It is unclear how my visual system can react to a secondary property though, as it seems its operation can be explained by physical properties alone. A possibility is that the red object stimulates my visual system just in virtue of its primary properties, and my experience of its redness is explained by how my visual system works. In any case, there must be a physical object out there for me to see.

As for the case of sound, if sounds are secondary events which are non-physical, it is equally unclear how it can stimulate my auditory system. We might wonder if the response used in the case of colour can be employed here. So, the suggestion would be that there is some physical event which stimulates my auditory system in virtue of its primary properties. It is, however, the way my auditory system responds which explains my experience's being as of a sound. Since the cause of an experience needs not to be identical to the entity represented in that experience, the sound needs not to be identified with that physical event.

This response does not work. If we accept it in the case of colour, then we should by parity conclude in the case of sound that my auditory experience is an experience of a physical event together with its secondary auditory properties. It would then be unexplained why we cannot identify that physical event as a sound, given that sounds are bearers of auditory properties. Or else if we accept this response in the case of sound as showing that sound are secondary events, then we should in the case of colour say instead that the red object is a secondary object which is not identical to the physical object stimulating my visual system. Similar things can be said for other sensory modalities, such that we can easily arrive at the absurd conclusion that everything we perceive, no matter via which modality, is a non-physical, secondary entity.

One possible response is to maintain the status of sounds as secondary entities by embracing the more traditional secondary property theory of sound. Probably we can classify it as a special version of dispositional property theory. Anyway, this does not seem to be a route Scruton would take. His view that sounds are pure events which happen to nothing is motivated by the idea that sounds can be detached from sound sources. Insofar as he still holds this latter idea, treating sounds as secondary properties would commit him to the consequence that secondary properties are free-floating properties which can have no bearer.

The existence of free-floating properties is even more implausible than the existence of pure events. We may, nonetheless, suspect that Scruton would be willing to pay this price, as he does not seem to be at any rate reluctant to accept the existence of pure events. I could, therefore, see no hope of persuading Scruton—if he were resurrected—to give up his view. I will now simply supplement the above discussion with a methodological consideration against any theory which posits secondary entities.

One observation which may motivate secondary property theory of colour, sound, smell, etc. is that no one-to-one relation between such properties and physical properties can be found. As a result, these properties should not be identical to the physical properties of their bearers. However, as I have said above, perceptual experiences are joint products of external entities and our perceptual systems. Suppose *p*, *q*, and *r* are primary properties. If all objects with either *p*, *q*, or *r* produce an experience of redness in a normal observer under normal conditions, then there are at least two possible ways to explain our experience of redness in terms of dispositions. Secondary property theory chooses to attribute to those objects the disposition to appear red to normal observes in normal situations. We may, however, instead attribute to a normal observer the disposition to have an experience of redness when being stimulated by those objects in normal situations. I see no reason to prefer the first way over the second one. Indeed, I shall try to show that the second way is more reasonable.

I think Scruton is right when he notes that "the science of secondary qualities is the science of perception" (2009, p. 56). However, instead of positing a secondary property shared by all entities which produce an experience of redness, a more reasonable development of this thought is that we should focus on the contribution from the perceptual system in the perceptual process. Although it looks easier to explain our experiences of redness in terms of secondary properties, redness, as Scruton admits, "denotes no explanatory property" (ibid.). By shifting our attention from the external entities to the internal workings of our perceptual system, a more promising direction of investigation is opened. More can be learnt if we bear in mind the fact that those "red" entities have very different physical properties, and then dig into our visual system and find out how it makes all those entities appear red in our experiences.

If positing a secondary property does not explain our experiences of redness, then there is no reason to add it to our ontology. Whatever colours are, we should not treat them as secondary properties. Similarly, taking sounds as secondary entities—no matter properties, objects, or events—cannot contribute to our understanding of auditory perception. In contrast, we should investigate how external stimuli and our auditory system jointly produce our auditory experiences, as well as other aspects of auditory perception. There is no reason to posit any secondary entity in explaining auditory perception, and so we should stick to a simpler ontology. Whatever sounds are, there are no secondary entities with which they can be identified.

6 Sounds as Compression Waves

This chapter is all about wave theory. The first four sections form the first half of this chapter and are general in character. By this I mean they do not assume any specific version of wave theory. I begin in §6.1 and §6.2 by putting forward two arguments for wave theory. The first one is an analogical argument which follows from my criticism of the analogical argument offered by property theorists. The second argument appeals to the explanatory power of wave theory. Next, §6.3 and §6.4 reply to two objections against wave theory which focus respectively on the spatial and temporal aspects of sounds.

I start detailing my version of wave theory in the second half of this chapter. First, in §6.5, I provide a metaphysical analysis of compression waves. Next, I examine what it is to hear compression waves in §6.6. These two sections form the backbone of my view. Then I reject another medial theory—wave pattern theory—in §6.7. After that, I move on to fill in more details of my view. The eight features of sounds discussed back in Chapter 3 are revisited in §6.8 to see how my view accommodates them. Lastly, §6.9 discusses four special sonic phenomena to further demonstrate the explanatory power of my view.

6.1 First Argument: An Analogy between Sounds and Light

The first argument for wave theory is not a self-standing one. It is embedded in the context where analogical arguments are used by property theorists to support their view. To the extent that this is a respectable move in the philosophy of sound, my first argument shows that a stronger analogical argument can be provided for wave theory. The conclusion we can arrive at can therefore only appeal to those who find the analogical argument for property theory plausible. In other words, my argument can only ground a comparative judgement favouring wave theory over property theory. Nonetheless, an interesting lesson can be learnt from my argument—we should avoid comparing sound sources with objects which merely reflect light in contemplating the similarities between the visual world and the auditory world.

I start by evaluating the most common way of analogising sounds and colours among property theorists. This reflection shows us a way to improve the analogy between sounds and colours. The improved analogy, however, is still not quite satisfactory. I then move on to show how wave theory can offer an even better analogy—one between sounds and light instead.

6.1.1 First Analogy between Sounds and Colours

We have seen a few property theories in Chapter 5. One thing they have in common is that their proponents tend to analogise sounds to colours. Pasnau (1999, p. 313) says that sounds are more like colours than odours. Just like we see colours as located at the bearers, we also hear sounds as located at their sources. Leddington (2014, p. 331) claims that sounds auditorily appear to be bound to the events producing them in the same way as colours visually appear to be bound to objects.

An obvious problem for these two suggestions is that the similarity between auditory and visual experiences does not necessarily imply that sounds and colours are also similar. In contrast, Kulvicki (2008a) does not appeal to phenomenology in drawing his analogy. Instead of comparing the subjective characters of visual and auditory experiences, he focuses on the objective conditions of our perception of these two kinds of properties.

As we saw in §5.1.2, Kulvicki's analogy is the most articulated one. Here is a quick reminder of his view. Just like colours of objects are their disposition to reflect light under illumination, sounds of objects are their dispositions to vibrate upon mechanical stimulation. Colours and sounds are perceived by means of light and compression waves respectively.

I mentioned two problems related to the two roles light plays in Kulvicki's account. First, light is the visual perceptual means. Since light has colour, if sounds are tightly analogous to colours, then it seems compression waves as the auditory perceptual means should have sounds as their properties. This implication is similar to Nudds's view that sounds as properties are patterns or structures of frequency components instantiated by compression waves (2009, 2010b). However, it is inconsistent with Kulvicki's view, because he locates sounds at sound sources rather than in the medium.

Second, light is *the* stimulant to reveal the colours of objects. In contrast, stimulants of sound sources do not form a natural kind. A "thwack", Kulvicki's term for a stimulant of a sound source, can refer to a compression wave or other kind of mechanical impact.

It might be objected that there are also multiple kinds of stimulant for colours. Heating a piece of iron can turn it red. Blue sparks can be seen when a Wint-O-Green lifesaver candy is crushed.[31] Such examples of light sources, however, show nothing for our purpose. Recall that Kulvicki specifically understands colours as dispositions of objects to reflect light. By definition, such dispositions can only be revealed by light. Colours of light sources and light cannot be understood along the same line, and hence they are irrelevant.

Table 1. First analogy between sounds and colours

Light Electromagnetic wave	STIMULANT IDENTIFIED WITH	Thwack Compression wave & other mechanical impact
	| stimulate ↓	
Reflective object	OBJECT	Sound source
Colour Disposition to reflect light	PROPERTY IDENTIFIED WITH	Sound Disposition to vibrate
	| produce ↓	
Light Electromagnetic wave	PERCEPTUAL MEANS IDENTIFIED WITH	Compression wave N/A
Colour	PROPERTY	No sound

Table 1 summarises Kulvicki's analogy and the two problems just discussed. Four disanalogies can be noticed. First, the stimulant for colours forms a natural kind, i.e. electromagnetic wave; but the stimulants for sounds include more than compression waves. Indeed, most often sound sources are stimulated by other kinds of mechanical impacts. Second, sound sources are the sources of compression waves, but reflective objects merely reflect pre-existing light. Third, light is electromagnetic waves, but compression waves are not identified with anything—at least not with sounds for distal property theory. Fourth, electromagnetic waves have colours, but compression waves do not have sound.

[31] Through a process called triboluminescence.

It seems the second disanalogy can also be found in Leddington (2014, 2019) and Pasnau (1999). Pasnau seems to only have reflective objects in mind in his discussion on colours. When he imagines a possible world "where colours typically last but a moment", he takes constant changes in surface properties or lighting conditions as the only two possible explanations of the short-livedness of colours (1999, p. 322). Neither explanation applies to light sources. Therefore, he compares sound sources to reflective objects only.[32] In contrast, Leddington uses as examples both light sources such as fireworks (2019, p. 625) and reflective objects such as tomatoes (2014, p. 332). However, this does not make a huge difference, as the disanalogy cannot be avoided unless he focuses exclusively on light sources.

This disanalogy could be avoided for those who accept the event view of colour. Accordingly, colours of reflective objects are properties of the events of absorbing and re-emitting photons which happen on reflective surfaces (Pasnau, 2009, p. 355). To the extent that such objects can be viewed as the sources of the re-emitted light, they are analogous to sound sources. Anyway, this leaves other disanalogies untouched.

As for the other disanalogies, the first one concerns Kulvicki's view only, so it need not be a problem for distal property theory in general. However, the third and fourth disanalogies follow from the basic idea of distal property theory that sounds are located at their sources but not in the medium, so any attempt to avoid them would amount to giving up the theory. As a result, distal property theorists can only try to refine their view by analogising sound sources to light sources instead. The idea would then be that sounds are properties of sound sources in the same way as colours are properties of light sources. None of the distal property theorists deny that light sources have colours, so we can expect that they would not resist revising their views in this direction. Let us consider if this suggestion works.

6.1.2 Second Analogy between Sounds and Colours

[32] Later in a different context, Pasnau (2009, pp. 361-363) notes that light sources are almost always ignored in the discussion of colour. However, while he acknowledges the difference between light sources and reflective objects, his aim is to provide a unified account of the colours of these two kinds of objects.

We attribute colours both to reflective objects and light sources but in different ways. For example, when the illumination changes from white to red, a white wall remains white, even though it looks red under the new illumination.[33] In contrast, if a light bulb changes from emitting white light to emitting red light, we would not say that it remains white. Rather, we would say that it was white, but now it is red.

It thus seems that the colour of a light source is closely connected to the colour of the light emitted. Perhaps what it is for a light bulb to be white is to emit white light. More generally, we may say that a light source inherits the colour of the light emitted by it.[34]

If sounds and colours are analogous in being properties of sound sources and light sources respectively, then sounds should also be properties of compression waves, such that sound sources can inherit the sounds of the compression waves they produce. However, this is not possible in distal property theory, because it locates sounds not in the medium. As a result, we can only draw the analogy shown in Table 2.

Table 2. Second analogy between sounds and colours

Light source	WAVE SOURCE	Sound source
Colour	PROPERTY	Sound
	| produce ↓	
Electromagnetic wave Light	WAVE IDENTIFIED WITH	Compression wave N/A
Colour	PROPERTY	No sound

There are three obvious disanalogies. First, once again, while light is electromagnetic waves, compression waves are not identified with anything in distal property theory. Second and relatedly, while light sources are sources of light, sound sources do not produce sound. Third, electromagnetic waves have colours, but compression waves have no sound.

[33] This is so in the standard reflectance physicalist account of colour but not in versions of the event view which refuse to accommodate colour constancy (Pasnau, 2009, p. 365).

[34] A similar but slightly different observation can be found in Pasnau (2009, pp. 361-362).

We have now seen that both the two analogies drawn between sounds and colours are not very tight. They could thus provide very limited support for distal property theory. We can also see that both analogies share a problem stemming from the refusal to locate sounds in the medium. We may, therefore, wonder if wave theory can offer a tighter analogy.

6.1.3 An Analogy between Sounds and Light

Let us begin by keeping the analogy between sound sources and light sources. To improve the second analogy above, we simply give up the idea of analogising sounds to colours. The first step is then to identify another property as the analogue of colours.

Recall that both light sources and light have colours and that the former inherit the colours of the latter. To keep the analogy tight, we need to identify a kind of property which can be attributed both to sound sources and compression waves, such that sound sources can inherit it from the compression waves produced. What properties do compression waves have? Wave theory has a suggestion: auditory properties such as pitch, loudness, and timbre.

Notice that the suggestion is not to analogise colours to *acoustic* properties such as frequency, amplitude, and spectrum. Although auditory properties are correlated with acoustic properties, the former cannot be attributed to sounds or sound sources independently of what it is like to experience them. In contrast, acoustic properties are mind-independent.

It is generally agreed that sounds have pitch, loudness, and timbre. Moreover, it is common to attribute these auditory properties to sound sources by reference to their sounds. The tuba is low-pitched because it plays notes in the bass range. The trumpet is loud because its sound is loud. The violin has a sweet timbre because its sound is sweet. It thus seems that sound sources inherit the auditory properties of their sounds in just the same way as light sources inherit the colour of their light. Wave theory simply makes this equivalent to saying that sound sources inherit the auditory properties of the compression waves produced. We can then draw the analogy in Table 3.

Table 3. An analogy between sounds and light

Light source	WAVE SOURCE	Sound source
Colour	PROPERTY	Auditory property
	| produce ↓	
Electromagnetic wave Light	WAVE IDENTIFIED WITH	Compression wave Sound
Colour	PROPERTY	Auditory property

This time, light, rather than colours, is the analogue of sounds. Just like light sources are sources of light, sound sources are sources of sound. All the problems of the two analogies in the previous subsections are avoided. To the extent that sounds are more analogous to light than to colours, people persuaded by the analogical argument for distal property theory should find wave theory more plausible.

This conclusion might not be able to help us defend wave theory against other theories of sound, as they, in general, do not rely on any analogy. Indeed, sometimes it is even stressed that sounds and auditory perception should not be modelled on their visual counterparts (e.g. O'Callaghan, 2007b). Theorists of this sort would therefore not be impressed by how tight the analogy between sounds and light is. It is for this reason that another attempt to defend wave theory is needed.

6.2 Second Argument: Emanation

Kalderon (2018a, p. 11) cites the following passage from Broad (1952) as motivating wave theory:

> But the noise is not literally heard as the occurrence of a certain sound-quality within a limited region remote from the percipient's body. It certainly is not heard as having any shape or size. It seems to be heard as *coming to* one from a certain direction, and it seems to be thought of as pervading with various degrees of intensity the whole of an indefinitely large region surrounding the centre from which it emanates. (p. 5, italics in the original)

It is suggested that such an "emanative phenomenology" naturally motivates wave theory. Although I am also trying to defend wave theory, I shall express my disagreement with Kalderon here. At least in my own case, I do not think that Broad captures accurately the phenomenology of our auditory experience. I instead agree with distal theorists that we do not hear sounds as coming from a certain direction like a flying arrow, nor do we hear them as pervading in a large region. However, while distal theorists think that it is because sounds are heard as located at the sound sources, my reason is that we do not hear sounds as being anywhere. We do hear something as located at the sound sources, but those are the sound sources themselves, and we can hear them correctly.

I will say more about what it is like to hear sound if wave theory is true in §6.6. For the moment, I just want to point out that it is a mistake to think that wave theory is motivated by the emanative phenomenology, because that is not what it is like to hear sounds. However, this does not pose any problem for wave theory. As a metaphysical theory of sound, wave theory need not be motivated by auditory phenomenology.

The usual reason to take auditory phenomenology into account is the thought that sounds are what we perceive, and the veridicality requirement that a theory of sound should not imply that auditory perception is massively erroneous about its objects. However, once we have realised that, in having an auditory experience, what people take to be sounds may instead be sound sources, it is unclear what the phenomenology of auditory experience can show for the metaphysics of sound. How we characterise the auditory phenomenology is informed by our conception of auditory perception. However, as I have tried to show in Chapter 2, we just do not understand our auditory experiences well enough to give an indisputable characterisation of the auditory phenomenology. An auditory experience can be characterised as representing a sound or a sound source as distally located. The first characterisation, combined with the veridicality requirement, would favour distal theories. But one man's modus ponens can be another man's modus tollens. Wave theorists, for example, can respond that the veridicality requirement just shows that we should characterise the experience in the second way.

Worse still, the applicability of the veridicality requirement to the metaphysics of sound is not entirely unquestionable. As Nudds (2015) forcefully argued, the function of hearing is to let us learn about the sound sources. The veridicality requirement would apply if we are investigating the nature of sound sources, but there is no direct reason why auditory experiences should represent sounds veridically if auditory perception is not evolved for hearing sounds.

I am not denying that hearing is completely irrelevant to what sounds are. I am simply deeply sceptical of how far focusing on auditory *experiences* from the subjective viewpoint can take us in the study of the metaphysics of sound. What seems to be a more feasible route is to direct our attention to auditory *perception* considered from the third-person viewpoint. This method allows us to gather facts of auditory perception, or *facts of hearing* for short, which can more neutrally constrain theories of sound. I believe that such facts constitute the strongest motivation and support for wave theory.

A note of clarification: auditory phenomenology and facts of hearing are not mutually exclusive. What it is like to hear one thing may constitute a fact of hearing that thing. What I want to highlight is that there are still other facts of hearing which have not received enough attention in the philosophy of sound.

Philosophers of sound in general emphasise the auditory phenomenology. This is obvious when we see that wave theory is rejected just because we do not hear sounds *as* propagating in the way compression waves do. I will respond to this objection in the next section. For the moment, I want to point out that although there is no emanative phenomenology, our conception of sound does consist of an idea related to the emanation of sounds: the idea that sounds are caused by sound sources. We examined this idea in §3.6 and we concluded that it cannot come from bi-model experiences of sound sources. Now I go further and propose that, like many discoveries in science, the fact that sounds are caused by sound sources is an unobservable fact which best explains certain observations.

As I said earlier, we do not hear sounds as propagating from their sources. There is no emanative phenomenology. Therefore, we cannot observe the generation and the propagation of sounds. Nonetheless, there are facts of hearing which allow us to infer that what immediately cause our auditory experiences are produced by and emanate from sound sources. Together with the thought that sounds are the immediate causes of our auditory experiences, it can be concluded that sounds are caused by sound sources and propagate into the surroundings. Further study reveals that compression waves are the only entities in the medium which have these two features. It is then natural and reasonable to identify sounds with compression waves, and thus we have wave theory.

What are those facts of hearing? There are many. We can identify differences in our auditory experiences which result from differences in the immediate causes but not from the event sources. For instance, when it is too noisy outside the house, I can simply close the windows without doing anything to the event source outside. That which is changed is the immediate cause of my auditory experience. Also, the closed windows change it by blocking its path into the house. Unless we take this causal intermediary as spatially extended, we cannot explain why closing the windows can have such an effect.

Besides, we can observe that our auditory experience of a faraway event is delayed for a short period by comparing it with the corresponding visual experience. This suggests that the causal intermediary is temporally extended as well. Moreover, the delay is proportional to our distance with the event source. The proportionality between the spatial and temporal extensions of the causal intermediary strongly suggests that it is something motion-like.

Apart from such facts of hearing, there is still another fact of sound which is related to the emanation of sound. It is the fact that we are not merely perceivers of sounds but also sound sources—we talk and speak to other people. This fact, in my opinion, has not received enough attention in the philosophical literature. What I have in mind is not the third-person observation that human beings are sound sources. This would be have great difference from the cases of other sound sources we hear. Rather, I am pointing to the fact that we are agents who make verbal sounds to address other people. When I speak, I hear my voice. I also witness the person I talk to responds to what I say. I am aware of myself as an agent who does things to another person *at a distance* using the sounds I make.

I also have experiences of what would make my voice less effective in addressing other people. If I talk to a person without facing her, she might not be able to hear what I say. If there is a glass wall in-between, we can see each other but my voice cannot get across to reach her ears. If I want to talk to someone upwind, my voice needs to be louder. All these observations suggest that when I talk to a person, I am producing something which can travel across the distance between her and me and interact with the environment. What is that thing? I speak; I make a sound. It is the sound I made. Sounds therefore are things which travel, or at least this is what we would conclude naturally. Of course, we do not hear sounds travel: there is no emanative phenomenology. Instead, we can see, hear, and act in response to the effects of their travelling in air.

From these numerous emanation-related facts, it is reasonable to conclude that sound sources and hearers are connected by certain entities which travel across the distance between them. This conclusion would not be disputed by other philosophers of sound. They might also concur with scientists in identifying that kind of entity with compression waves. However, the further step of identifying sounds with compression waves is what those philosophers reject. The most plausible alternative is that compression waves are not sounds but merely what make sounds audible from a distance. Accordingly, the observations solicited above merely suggest that properties of compression waves account for the appearance of sounds in our auditory experiences. The sounds themselves, as well as the sound sources, are not affected when we manipulate the compression waves.

I admit that this alternative would be empirically equivalent to wave theory at least in ordinary cases. Therefore, a likely method to choose between them would appeal to purely theoretical considerations. As far as I can tell, the alternative account is only proposed as a supplement to distal theory to leave a role for compression waves to play in determining our auditory experiences given that sounds are located at their sources. So, if we do not accept distal theory, then the alternative account is not motivated.

I have examined all the existing versions of distal theory in Chapter 5 and showed that they are all unsatisfactory. At the more general level, distal theorists take the apparent distal locatedness of direct auditory objects as a crucial desideratum, and they simply take such distally located objects as sounds without further justification. In other words, their theory is based on a questionable characterisation of our auditory phenomenology. Once the thought that sounds appear to be distally located is discarded, there is hardly any direct evidence for distal theory.

I mentioned earlier that apart from auditory phenomenology, there are still other facts of hearing. Now we can see that distal theorists have put too much weight on the former, while at the same time treating the latter merely as facts about the audibility of sounds. I take myself as having shown why this makes distal theory ill-founded.

The remaining option for distal theorists is to show that their theory is better than other theories—including wave theory. Their objections against wave theory are therefore crucial to their project. I shall then proceed to respond to them.

6.3 First Objection: Spatial Misrepresentation

The most common objection against wave theory, which I call the spatial misrepresentation objection, concerns the spatial representation of sounds.

> S1. Compression waves propagate from their sources in the surrounding medium.
>
> S2. We never hear sounds as moving in the way compression waves propagate.
>
> S3. Therefore, sounds are not compression waves.
> (i.e. Wave theory is false)

S1 should be an indisputable scientific fact. S2 also seems to be quite plausible. I have never heard anything as moving like a giant balloon inflating rapidly from a sound source, or like an arrow flying from a sound source to me.

It is obvious that the veridicality of our auditory experiences of sounds is assumed, and this is, as I said previously, a questionable assumption. Nonetheless, I will accept it for the moment and see if we can satisfactorily respond to the objection.

If wave theorists accept both premises, then they can only challenge the inference. Let us consider two possible ways to do so.

6.3.1 First Response: Not Hearing Sounds

There is one radical response, that is, we deny that we hear sounds. If we do not hear sounds, then, of course, we do not hear them as propagating in the way compression waves do. But then S2 will become irrelevant to what sounds are, and hence S3 does not follow.

This idea may seem too radical, but perhaps wave theorists may motivate it by considering a parallel argument against the identity between light and electromagnetic waves:

L1. Electromagnetic waves propagate from light sources to the surroundings.

L2. We never see light as moving in the way electromagnetic waves propagate.

L3. Therefore, light is not electromagnetic waves.

To be consistent, wave theorists should reject this argument, and I believe many people would be reluctant to accept L3.

In this case, if we challenge the inference by denying that we see light, the consequence does not seem to be very devastating. Hilbert (2005, pp. 150-151), for instance, holds on phenomenological grounds that we see how an object is illuminated but not the illumination itself. Perhaps all we see are objects including light sources and reflective objects. Such objects might look bright, but we do not see light in addition to seeing them. Light as electromagnetic waves can cause visual experiences, but it is not itself experienced.

So, wave theorists might suggest that all we hear are objects such as sound sources. Sounds as compression waves are causes of auditory experiences, but they are not themselves experienced.

Is there any other reason for wave theorists to deny that sounds are experienced? The eliminativism of Young (2016) is one suggestion. Recall that the weaker interpretation of his view is to deny the role of sounds merely in auditory perception. Accordingly, no matter what sounds are, we need not to incur them in explaining our auditory experiences. Therefore, we should simplify our theory of auditory perception by denying that sounds are experienced.

Young's view is based on a theoretical consideration, that is, the simplicity of his theory of auditory experience. It thus relies on the presumption that his theory is correct. Strictly speaking, wave theory is consistent with Young's view, since it is possible that sounds are compression waves but do not figure in our auditory experiences. Nonetheless, wave theorists probably would not want to add a further commitment to this theory of auditory experience.

Besides, I do not think that a description of auditory experiences can be complete without mentioning compression waves. As I have argued in §4.2.1, Young's sound-less account cannot explain some aspect of our auditory experiences. The example I used is our echo experiences, a more detailed discussion of which can be found in §5.2.6 in connection with distal event theories in general.

Here is a quick review. If wave theory is true, not hearing sounds means not hearing compression waves. However, when I hear an echo one second after I hear an explosion, the most reasonable thing to say is that I hear the delay between the arrival of the primary wave and the arrival of the reflected wave. This experience is impossible unless it can represent compression waves. Therefore, wave theorists cannot accept a sound-less account of our echo experiences, and hence they cannot employ the first response that we do not hear sounds.

6.3.2 Second Response: Not Hearing the Propagation of Sounds

The more plausible way to challenge the inference in the spatial misrepresentation objection is to say that although sounds as compression waves are experienced, we fail to hear their propagation. Let us consider this argument:

W1. Water flows in the river.

W2. If we look at the river from afar, we cannot see anything as moving in the way water flows.

W3. Therefore, if we look at the river from afar, we cannot see the water in the river.

This argument obviously fails because it is more reasonable to say that we simply fail to see the flow of the water. Although this argument is not exactly parallel to the spatial misrepresentation objection, it seems the latter can also only warrant the conclusion that we fail to hear the propagation of compression waves.

I think this is the best response to the spatial misrepresentation objection. Wave theorists should hold that whatever it is like to hear compression waves, it is not to hear them as propagating. This negative claim immediately prompts us to ask for a positive characterisation of what it is like to hear sounds. Indeed, the lack of such a positive characterisation leaves open a way to strengthen the spatial misrepresentation objection, which shall be examined before we move on.

6.3.3 Not Hearing Sounds as Distally Located

Here is the strengthened version of the spatial misrepresentation objection:

S1. Compression waves propagate from sound sources to the surroundings.

S2′. We hear sounds as located at where their sources are.

S3. Therefore, sounds are not compression waves.

S2′ positively asserts what it is like to hear sounds, that is, to hear them as located at where their sources are. My response to this strengthened objection is that S2′ is an inaccurate characterisation of our auditory experiences.

There are two possible interpretations of S2′. The first interpretation says that we hear a sound as sharing the same location with its source. We hear the relation of co-location. To do so, both the sound and its source should appear in my auditory experience, and they need to appear at the same location. However, this does not seem to be phenomenologically accurate.

Suppose I close my eyes and tap the table. This event causes me to have an auditory experience. Within this experience, there appears to be only *one* individual at a certain distance away from me. According to the identification problem of auditory objects, it may be the event of my tapping the table or the sound of this event. However, whatever the auditory object is, there does not appear to be two things sharing the same location.

The second interpretation says that we simply hear a sound as located at a certain location. As a matter of fact, this location is the location of the sound source, but this fact does not enter the content of my auditory experience. Instead, we learn it in some other ways.

This interpretation fits better with the phenomenal character of my experience, as it requires only one thing to appear in my experience, namely, the sound. But then it seems we have changed the question of what it is like to hear a sound to the question of what it is like to hear a sound source. If only one thing appears in my experience, and it is a sound, then it seems sound sources are not experienced. However, in Chapter 2, we have seen no good reason to think that that thing we hear is not the sound source itself, and hence this does not show that what we hear is a sound rather than the sound source.

Since the two interpretations of S2′ are both implausible, we can reject the strengthened objection by rejecting S2′. In contrast, S2 in the original objection seems to be a more accurate characterisation of our auditory phenomenology. That said, we might wonder why S2′ would seem to be plausible for the opponents of wave theory. The main reason, I suppose, is that our auditory experiences do represent some entities as located at where the sound sources are. If S2′ is implausible, then wave theorists should tell us what those entities are. The answer is simple: those are the sound sources themselves, and our auditory experiences correctly represent them as being at where they are. This answer, of course, leaves unexplained what it is like to hear sounds. More specifically, it says nothing about whether and how sounds are represented spatially. I return to this issue in §6.6.

6.4 Second Objection: Temporal Misrepresentation

The second objection against wave theory appeals to the fact that a wave has a lifetime longer than what we take the duration of a sound to be (O'Callaghan, 2007b, pp. 43-46). Consider this actual case (Heller, 2013, p. 567). On 21st September 1921, a silo exploded in Oppau, Germany. The explosion was heard more than 300 km away. Since compression waves refract in the air due to atmospheric conditions such as the presence of wind and the temperature gradient across altitudes, it is not known how exactly the wave travelled from the explosion. However, even if we assume that the wave travelled along a straight path at 343m/s (the speed of compression waves in 20°C air at sea level), it still took around 15 minutes for the wave to travel 300 km, a time much longer than the duration of the explosion.

How long did the sound of the explosion last? Ordinarily, we would say that it lasted for roughly how long the explosion lasted, i.e. just a brief instant. This is suggested by our own experience, which also lasted for the same duration. On the other hand, if wave theory is true, it seems the answer should rather be that it lasted until the compression wave completely died out in the atmosphere hundreds of kilometres away, i.e. at least more than 15 minutes. The objection is then that wave theory implies that the length we ordinarily experience a sound as having is not its real length. We have massive and systematic illusions about the lengths of sounds.

O'Callaghan suggests that what we experience is not the duration of the compression wave itself but our encounter with it. When the wavefront of a compression wave reaches me, I start hearing a sound and take it as beginning at this moment. When I encounter the far spatial boundary of that compression wave, we stop hearing the sound and take that moment as its end. If wave theory is true, what we ordinarily do is to mistakenly treat the duration of our experience as the lifetime of the compression wave itself.

This objection is basically the temporal version of the spatial misrepresentation objection. Indeed, wave theorists can reply similarly. The so-called "lifetime" of a compression wave is the duration of its propagation. Since I hold that we fail to hear the propagation of compression waves in §6.3.2, I should also hold that we fail to hear the "lifetimes" of compression waves. As for the thing which is experienced as lasting for only a brief instant, it is the explosion itself. Since the explosion is experienced as lasting for the duration it has, there is no illusion.

This response and my response to the spatial misrepresentation object would be more phenomenologically plausible if they are considered together. An auditory experience does not represent one distally located object alongside with one temporally extended object; rather, it represents an object which is distally located *and* temporally extended. It is thus more plausible to identify this auditory object with a single entity which has the right spatial *and* temporal features. The sound source—more precisely, the event source—which causes the auditory experience in question obviously meets this criterion, and it is unclear what else could meet the same criterion equally well. Therefore, responding to the two misrepresentation objections by identifying the distally located and temporally extended auditory object with the event source is the most reasonable response.

Again, it is still unexplained what it is like to hear sounds. Holding that the distally located and temporally extended auditory objects are sound sources does not imply that sounds are not represented temporally in auditory experiences. I return to the experience of sounds in §6.6.

For now, it is worth pointing out that both the spatial misrepresentation objection and the temporal misrepresentation objection focus on how compression waves are represented in auditory experiences if wave theory is true. They do not directly challenge the identification of sounds with compression waves, but rather reject the alleged implications of wave theory on our auditory experiences. It is, however, questionable whether wave theory really has those implications. Indeed, what I have tried to do is to show that this is not the case. For example, in responding to the original version of the spatial misrepresentation objection, I accept both premises but conclude that we simply fail to hear the propagation of sounds. This is in effect a rejection of the hidden assumption that wave theory implies that we should hear the propagation of sounds. There is no *a priori* reason to assume this.

I agree that wave theory constrains our proper understanding of auditory perception. However, compression waves are very different from the ordinary material objects we are more familiar with. Therefore, we should proceed slowly in determining exactly what implication on auditory perception can be drawn from wave theory. This should be done by first making sure what compression waves are.

6.5 A Metaphysics of Compression Waves

We should not think of a compression wave as an inflating balloon or a flying arrow. There is no object travelling from a sound source to an observer and beyond. Consider the case of airborne sounds. In reality, there are only air molecules bouncing back and forth in the air. What happens when an event source produces a compression wave is that the thing source displaces air molecules immediately around it in an orderly way.

Event sources in my view should, therefore, be the disturbance events identified with sounds in disturbance event theory. They are in turn individuated in terms of the events causing them. Drawing a bow on a string of a violin drives the top plate of the instrument to vibrate continuously. Tapping on the top plate causes it to vibrate but only instantly. If a player taps on the top plate when she is drawing the bow across the E string, two disturbance events are happening at the boundary between the instrument and the surrounding medium: one caused by the drawing and one caused by the tapping. Therefore, there are two event sources in my view.

Other wave theorists may instead identify event sources with some larger events which include such disturbance events as proper parts. For ease of expression, sometimes I may use the events which produce event sources to stand for those event sources. For instance, when I simply refer to a violin performance as an event source, I should be understood as talking about the disturbance event which happens at the boundary of the violin.

An event source produces local fluctuations of air pressure at every location within the surrounding region. The pressure difference between the affected region and the region further away causes the air molecules in the latter region to change in the same way. The pressure fluctuation is thus propagated into regions further away from the event source. This process continues until all the acoustic energy runs out. The propagation of pressure fluctuation should, therefore, be understood as a causal event rather than a kind of locomotion. Since the medium is continuous, the propagation of local pressure fluctuations does not have discrete steps. This gives rise to the impression of something travelling across a distance in typical visual illustrations of compression waves.

A compression wave is no more than the propagation of pressure fluctuation in the medium, and hence it is a causal event. It can be described in different ways. As a whole, it is an event which happens to a certain volume of an elastic medium during a certain period. It is therefore both temporally and spatially extended. We can talk about when and where it begins, when and where it ends, the region it spreads through, and its lifetime. The propagation begins from the sound source and ends when all the acoustic energy from that source runs out. These marks the lifetime of the compression wave. During this period, the compression wave spreads through the region between the sound source and the furthest locations the pressure fluctuations of which are caused by the event source.

We can also describe how the propagation happens by holding either the spatial dimension or the temporal dimension fixed. At a fixed location, there is a period in which the fluctuation of air pressure there is caused by the event source. During that period, the propagation happens to the medium at that location and can be described in temporal terms, such as frequency, spectrum, and attack and decay patterns. The pressure fluctuation at that location can in itself be counted as an event which is also a spatial part of the compression wave. Let us call such a local event "constituent fluctuation" to highlight its status as a local pressure fluctuation which constitutes the compression wave. Features of a constituent fluctuation can then be counted as the local features of the compression wave at that location. This means that if a feature needs to be described in terms of a spatial extension, it is not a local feature.

Alternatively, we can focus on an instant and single out the region the air pressure of which is determined by the source of the compression wave. We can then say that the propagation event is happening throughout that region at that moment and describe it in spatial terms, such as wavelength and attenuation. These are the instantaneous features of the compression wave at that moment. If a feature needs to be described in terms of a temporal extension, it is not an instantaneous feature.

What I have just said concerns the simple case where only one compression wave is present. In the more common situation, multiple compression waves coexist in the same environment. Compression waves can overlap with each other. The superposition of compression waves can be described as multiple propagation events happening at the same location within the same period, or as multiple propagation events happening at the same instant over the same region of space. The first way is the more common one, and I will focus on it. However, what I say in the following applies *mutatis mutandis* to the overlapping compression waves when they are described in the second way.

When compression waves overlap, the local pressure fluctuations in the overlapping regions are the combined effects of the overlapping compression waves. Let us call such a combined effect "total fluctuation" to distinguish it from constituent fluctuations. A total fluctuation at a location is then equivalent to all the constituent fluctuations there combined. When the distinction between total and constituent fluctuations is not important, I will simply use "local fluctuation" to refer to the pressure fluctuation at a fixed location.

In terms of metaphysics, there is always a determinate answer as to how a total fluctuation should be analysed into constituent fluctuations. This is not determined in qualitative terms, as constituent fluctuations of the same compression wave can have very different qualities at different locations. Amplitude decreases with distance. Spectral composition also changes with distance as the amplitude of higher frequency components decreases at a higher rate. Therefore, there is no single set of acoustic properties which can serve as the criteria of identification.

Moreover, a total fluctuation considered in itself does not contain any internal structure among its frequency components. Frequency components of any frequency at any amplitude can be mixed at the same location, and there is no empirical difference between the absence of a frequency component and the presence of that component at zero amplitude. Therefore, it is possible to say that every total fluctuation contains the same number of frequency components—*viz.* all—but at different amplitudes. Moreover, when we focus on one frequency, it is always possible to divide it into infinite components at the same frequency with infinitesimal amplitude.

The only way in which total fluctuations can be analysed into component fluctuations is to appeal to the nature of compression waves. A compression wave is a causal event which connects the pressure fluctuation at every location within its reach to its event source. A total fluctuation at a location would be different if one of the event sources were absent. The difference is then the causal effect of that event source at that location, and hence can be distinguished as the constituent fluctuation of the compression wave produced by that event source. Since all the constituent fluctuations of the same compression wave can be identified in this way, this implies that the entire compression wave can be identified by its event source. The event source thus determines the identity of the compression wave it produces.

By identifying sounds with compression waves, wave theory is, therefore, treating sounds as propagation events of pressure fluctuations. A sound can be described as a whole in terms of when and where it begins and when and where it ends. It can also be described locally. Doing so would be equivalent to describing the constituent fluctuation at the specified location. A surprising implication of wave theory is that sounds can also be described instantaneously in spatial terms. For instance, it has a certain volume at a moment in virtue of happening to a certain volume of the medium. Such an instantaneous spatial property of a sound is important in determining which perceivers in the surroundings can hear it.

With this metaphysics of compression waves in hand, we can now move on to discuss our perception of sounds.

6.6 Hearing Compression Waves

We hear sounds. For wave theorists, this means we hear compression waves. Despite its simple appearance, this claim can be separated into two different claims. The first one is the causal claim that compression waves cause our auditory experiences; the second one is the representational claim that compression waves are represented in auditory experiences. These are two independent claims, and wave theorists need not accept both. In this section, I explore these two claims in light of the metaphysics of compression waves presented above.

There are two crucial problems which need to be solved: the identification problem of auditory objects and the unification problem. In Chapter 2, I have already argued that it is more plausible to identify the distally located and temporally extended auditory object with the event source. As I claimed in §6.3.3 and §6.4, this is compatible with the possibility that sounds are also represented in auditory experiences. The remaining problem—i.e. the unification problem—is then to explain how a sound and its event source can be represented at the same time without there appearing to be two individuals in the auditory experience.

Different theories of sound have different solutions to the unification problem. For example, property theory can say that an auditory experience represents a sound source—thing source or event source—as instantiating a sound, and it is no mystery that an object or an event appears with its properties in our perceptual experience. As for distal event theory, it can say that when we hear a sound, we are hearing a thing source undergoing a sounding event. Our auditory experience thus simultaneously represents both the sound and its source.

In contrast, no simple answer is available for wave theory. Compression waves are entities in the medium whose subsistence does not depend on their sources. They and their sources occupy different spatial locations and have different persistence conditions. It is therefore very puzzling how they can be unified into one individual in an auditory experience.

I argue that although compression waves are causes of our auditory experiences, they are not represented in their entirety. Instead, our experiences represent the constituent fluctuations at our locations. In other words, we can only hear in the representational sense a local part of a compression wave. This latter representational fact, however, is a fact of hearing learnt only through reflection. In terms of phenomenology, hearing a constituent fluctuation is just like hearing its event source.

6.6.1 Hearing in the Causal Sense

Consider the causal sense of hearing first. It should be indisputable that we hear total fluctuations in this sense because our eardrums are set into motion by total fluctuations. Since the total fluctuation at a location is simply the combination of all constituent fluctuations at that location, this means we also hear constituent fluctuations in the causal sense. Nonetheless, the constituent fluctuations we hear do not cause our auditory experiences individually but as a mixture. This can be seen in the case of complete destructive interference, where two or more constituent fluctuations completely cancel each other and result in zero total fluctuation. Although there are constituent fluctuations, the absence of total fluctuation has no causal power to generate any auditory experience in the perceiver. Therefore, while it is not wrong to say that we hear constituent fluctuations in the causal sense, it is to this extent more appropriate to say instead that we hear total fluctuations in the causal sense.

We hear only those two total fluctuations directly next to our eardrums at the moment of hearing them.[35] This implies that we hear the compression waves which are happening at those two locations but only in a restricted sense: we do not hear the propagation of pressure fluctuation. At every moment, only the total fluctuations in direct contact with our eardrums are heard. Their causes as well as the causation between them are inaccessible. Therefore, we can only hear the local features of compression waves, without also hearing the global features concerning how the pressure fluctuation is propagated and the instantaneous features which spread through an extended region. This explains my claim in §6.3.2 that we fail to hear the propagation of compression waves.

It might be suggested that binaural hearing provides the opportunity of hearing a constituent fluctuation on one side causes another constituent fluctuation on the other side, such that the propagation can be heard. There are two problems with this suggestion. First, it restricts the possibility of hearing compression waves *qua* propagation events to creatures with two functioning ears, but it is implausible to deny this possibility only to those who are deaf on one side. Second, the two constituent fluctuations at the two ears are not causally connected in the right way. The causal path does not make a round trip to one eardrum and then move on to another eardrum. Instead, two separate causal paths link the two constituent fluctuations back to the event source. Therefore, no propagation of pressure fluctuation from one ear to another ear is there for us to hear.

6.6.2 Hearing in the Representational Sense

As for hearing in the representational sense, the case is more complicated. In an ordinary auditory experience, nothing appears to be at our eardrums. Phenomenologically speaking, we hear only something at a distance from us, and this thing is the event source. It is therefore far from obvious that the constituent fluctuations causing our auditory experiences are represented. However, there are occasions where we can identify in our auditory experience phenomenal features which do not correspond to the physical features of the event source.

[35] I glossed over the ~0.05 seconds processing time in the brain (Nevid, 2016, p. 100).

Consider this case. Suppose an orchestra is playing on the stage and you are allowed to walk around in the auditorium to choose the seat with the best sound. When you walk around, you hear the orchestra as staying at the same place, while at the same time you are aware of how your auditory experience varies systematically with your changing location. You can thus realise that your experience represents some features distinct from those of the performance, and it thereby also represents the bearer of those features. Since these features vary with your location, they are tied to your location rather than to the performance. However, these features do not appear to be at your location, as nothing appears to be so located. Rather, their location can only be learnt through reflecting on your experience. My suggestion is that such features are features of the constituent fluctuations next to your eardrums, and hence they are local features of the compression wave produced by the performance.

Your auditory experience in the above situation is both an experience of the performance and an experience of the compression wave. It is an experience in virtue of which two facts of representation obtain: the fact that it represents a performance and the fact that it represents a compression wave. The second fact should not be taken as having any implication on what it is like to hear compression waves. Indeed, at least in ordinary cases, auditory experiences always appear to be about event sources. Even in the above example of walking around in the auditorium, it would not be phenomenologically accurate to say that you hear something distinct from the performance. There is one sense in which *the performance* appears softer when you walk over to the back of the auditorium; there is another sense in which *the performance* appears equally loud despite your changing location. We may call the first "perspectival loudness" and the second "intrinsic loudness". Phenomenologically speaking, both perspectival loudness and intrinsic loudness appear to be features of the performance. Neither of them appears to belong to something distinct from the performance. It is only through reflection that we realise that the perspectival loudness should not be attributed to the performance but rather be treated as the local loudness of the compression wave.

Notice that in the representational sense of hearing, we hear constituent fluctuations rather than total fluctuations. This is because our auditory system analyses the detected total fluctuations into smaller units which in ideal cases correspond to the constituent fluctuations contained in the total fluctuations.

A likely alternative account is to take the perspectival loudness as really a property of the performance on the stage. There is no need to deny that the two constituent fluctuations next to your eardrums immediately cause and thus explain the phenomenal quality of your auditory experience. However, this does not imply that your experience represents such local events—a representation need not represent its causes.

It can be expected that distal theorists would prefer this alternative explanation. This is because they can avoid turning sounds—distal entities at the sound sources for them—into indirect objects of our auditory experiences. Even if they reject sonicism, it is unlikely that they would accept that sounds as distal entities are heard in the representational sense via hearing something in the medium. However, since I am now merely explicating what it means to hear compression waves in my wave theory, I need not reject this alternative explanation here.

Since in terms of phenomenology it seems we only hear event sources, there is a straightforward answer to the question of what it is like to hear compression waves: it is like hearing event sources. It is thus tempting to describe compression waves as perceptually "transparent": if you attempt to attend to a compression wave, you will end up attending to the event source and its features. However, there is a worry: if compression waves are not something we can single out in auditory experiences, it is then quite perplexing how auditory experiences can in any sense be counted as experiences of compression waves. Since sounds are compression waves in wave theory, this threatens the status of sounds as objects of auditory experiences.

We may respond by questioning the accuracy of describing compression waves as perceptually transparent. Consider the case of visually transparent objects. The pane of glass I see through does not share the visual properties of the object behind it. The object is, say, red, but the glass is colourless. Perhaps the object is a bright light source, but the glass is not bright. In the right situation, the pane of glass is invisible. However, this does not seem to be the case for compression waves. As I said in §6.1.3, sound sources inherit the auditory properties of their sounds. Also, saying that compression waves are inaudible may not sound right even to those who deny that we experience them.

The case of compression waves appears to be more like the case of photos, where the representations and the represented objects share visual properties. As Walton (1984, p. 252) says, photos are transparent but not invisible: we see them while seeing objects in them. Shall we then understand experiences of compression waves as analogous to experiences of photos? Doing so would then be treating compression waves as representations of their sources. There is a serious difficulty for this suggestion. It seems our ability to see objects in photos is at least partly explained by our ability to see those objects in real life without the causal mediation of photos. However, it is impossible to hear sound sources without the causal mediation of compression waves. It is therefore unexplained how we can hear event sources through hearing compression waves. Moreover, seeing an object in a photo is an indirect way of seeing it, and the indirectness is explicit in the experience. In contrast, hearing sound sources is not explicitly indirect—it is phenomenologically direct.

Visually transparent objects—visible or not—do not seem to be a good model for us to conceive of our experiences of compression waves. Therefore, it would probably be better to understand experiences of compression waves in some other sense. The analogy between sound and light introduced in §6.3.3 offers a new suggestion. Since sounds and sound sources are analogous to light and light sources, perhaps we can understand experiences of sounds and their sources by comparing them with experiences of light and light sources.

Philosophers typically deny that we visually experience light when we look at reflective objects. Hilbert (2005, pp. 150-151), for example, says that we see how an object is illuminated but not the illumination itself. However, the case of looking directly at a light source is sometimes treated as an exception to this claim (e.g. Casati, 2018, p. 167). Let us then accept that light is visible at least when we look at a light source. Do we see light as something distinct from the light source? Phenomenologically speaking, this is not the case. Rather, the light seems to be inseparable from the light source. It does not look like properties such as colour in appearing to be bound to an object. Nor does it appear to occupy a location different from that of the light source. Indeed, there does not seem to be any difference between seeing the light and seeing the light source.

The case is like what I have said about hearing a sound and hearing a sound source. The identification problem of auditory objects introduced back in §2.1.3 may arise from the experiential indistinguishability between hearing a sound and hearing a sound source. I propose that our experiences of waves and wave sources should be treated as a *sui generis* kind of perceptual phenomena which should be studied in their own right. For the moment, I can only discuss an observation and provide what I take to be a plausible preliminary explanation.

I shall now focus back on the case of sounds and sound sources, but I intend my claims to be transferable to the case of light and light sources. A crucial idea is that compression waves are the perceptual means by which we hear event sources. A compression wave carries information about its event source in virtue of the systematic relation between its local acoustic properties and the acoustic properties of the event source. Our auditory system takes advantage of this relation to retrieve information about the event source. This information retrieval process operates in a way which generates auditory experiences with certain auditory phenomenal properties. Such phenomenal properties represent both the local acoustic properties of compression waves and the intrinsic acoustic properties of event sources.

This qualitative aspect does not exhaust the content of auditory experiences: the locations of event sources are also represented. When interpreted with certain assumptions by our auditory system, constituent fluctuations can provide information about the distance between an event source and a perceiver in virtue of their spectra. To locate an event source, however, the auditory system still needs to determine the direction from which the compression wave comes.

Directional hearing partly relies on information provided by the interaural time difference (ITD) and the interaural level difference (ILD) between the stimuli of the two ears. Both ITD and ILD are results of the two different paths linking the two ears to the sound source. The two paths may or may not have equal length, depending on the location of the sound source relative to the perceiver on the horizontal plane. A compression wave enters the ear closer to the sound source first, resulting in an ITD between the arrivals of the compression wave at the two ears. Also, on the path between the sound source and the ear further away from it, there is an obstacle—the head. The compression wave thus needs to diffract around it before getting into that further ear, and this leads to a drop in the amplitude and thereby an ILD between the constituent fluctuations at the two eardrums.

Strictly speaking, ITD and ILD are not properties of the constituent fluctuations but quantities generated in the hearing process: they are relational properties between the two stimulation events rather than properties of the stimulants themselves.

In the resulting auditory experiences, such locational information is integrated with the qualitative information. Indeed, in ordinary auditory experiences, the qualitative aspect and the locational aspect seem to be equally salient. The auditory object thus appears to be an individual which has certain auditory properties and is located at a certain distal location. Since this auditory object appears to have both the qualitative and the locational aspects, it is more phenomenologically accurate to characterise an auditory experience as like hearing an event source rather than constituent fluctuations which do not have the represented spatial features.

We can now see a disparity between the spatiality and temporality of how compression waves are represented. Compression waves can be temporally represented because their local features at the eardrums are accessible by the auditory system. The same phenomenal properties of an auditory experience, such as phenomenal duration, pitch, loudness, timbre, etc., represent both the perspectival duration, frequency, amplitude, and spectrum of the event source and the local duration, frequency, amplitude, and spectrum of the compression wave.

In contrast, compression waves are not spatially represented. Again, our perceptual access to compression waves are limited to their local features at the eardrums. The auditory system has no mechanism to pick out the locational information about the compression waves. The only locational information retrievable from the stimulants concerns the location of the event source instead. Therefore, compression waves are only temporally but not spatially represented.

In ordinary cases, the qualitative and the locational aspects seem to be equally salient in our auditory experiences, and hence event sources are the foci of our auditory attention. This does not mean that we cannot focus on compression waves if we want. When I attend to the perspectival auditory properties of an event source in my auditory experience, I implicitly also attend to the local auditory properties of its compression wave at my eardrums, because they are represented by the same phenomenal auditory properties to which I am attending. However, since the phenomenal location only represents the location of the event source, I can attend specifically to the event source if I focus on its location. In contrast, I suspect that it is impossible to attend only to the compression wave, because our auditory experiences do not have any phenomenal properties which represent features of compression waves only.

One minor worry here is that there are two constituent fluctuations of the same compression wave causing an auditory experience, but they are not experienced separately. It might then be questioned whether they are really experienced at all. I believe the most plausible reply to this worry is that the resolution of our auditory experiences is not high enough to represent the difference between the two local fluctuations. Our auditory system is sensitive to the tiny differences and makes use of them, for example, in the localisation of sound sources. However, such differences are not reflected in the auditory properties experienced. In most situations, hearing something with one ear covered would not have any noticeable difference in auditory properties from hearing the same thing with the other ear covered. That said, I should note that in some cases, the difference becomes noticeable if it is large enough. The most obvious case is hearing music with only one earphone on. It is obvious in such an experience that the music and its qualities are represented only on one side. To this extent, we can hear constituent fluctuations individually.

My solution to the unification problem is that, although there seems to be only one thing present in my auditory experience, this experience actually represents both the local parts of a compression wave and its event source. A sound, identified as a compression wave, is experienced not in its entity but only partially. However, we cannot experience the most essential feature of it—i.e. the propagation of pressure fluctuation. There is no phenomenological difference between experiencing a sound and experiencing its source, since they are indeed the same act of auditory experiencing. Hearing a sound is just like hearing its event source. Equivalently, hearing an event source is just like hearing its sound. However, we can realise through reflection that two facts of hearing, or we may say, two facts of auditory experience, obtain in virtue of this experience: the fact that it represents two constituent fluctuations and the fact that it represents an event source.

6.6.3 Sonicism

In Chapter 2, I argued that sonicism cannot serve as a theoretically neutral ground for any theory of sound. Nonetheless, it may still turn out to be true. I proceed to examine it in light of my wave theory in this subsection.

When we focus on the case of sound sources, sonicism is obviously true in the causal sense. Sounds are compression waves, and our ears are immediately stimulated by local parts of compression waves. Therefore, sound sources can produce auditory experiences only via the causal mediation of compression waves. It is, however, unclear whether this can be generalised to all non-sound entities.

Consider the case of bone conduction. Bone conduction involves multiple mechanisms, one of which is of particular interest to us here: inertial bone conduction (Gelfand, 2018, pp. 79-80). The human skull responds to the vibration directly applied to it in different ways depending on the frequency of the vibration. At lower frequency (e.g. 200 Hz), it moves as a whole—i.e. the entire skull moves in the same direction. However, the ossicles are suspended in the middle ear like pendulums. Their inertia stops them from moving with the skull and results in relative movements between them. Such movements are basically the same as those which follow the vibration of the eardrum in the case of air conduction, and so the inner ear is stimulated in the same way.

What makes this case interesting is that the whole process of inertial bone conduction can involve no compression wave. The real human skull can be compressed, but even if it is completely rigid, nothing essential would be affected in the process of this kind of bone conduction. So, if a vibrator is applied directly to the side of such a completely rigid skull, nothing is compressed and hence this is no compression wave. According to my view, there is no sound, although the subject can hear the signal generated by the vibrator. Therefore, the resulting auditory experience is not caused by the vibrator (a non-sound entity) via a sound.

The vibrator is, in fact, a sound source, as its vibration in air generates a compression wave. However, in the current case, it is not heard in virtue of its sound, so it is not heard *qua* sound source. If the vibrator is not surrounded by an elastic medium, it cannot produce any compression wave and hence it is not a sound source. Nonetheless, a subject can still hear it through bone conduction if it is applied to her head directly.

I do not know what sonicists would say regarding inertial bone conduction. One option is to say that although the vibration of the vibrator produces an auditory experience, this is not the right kind of causal process to qualify the vibrator as being heard. The experience is, accordingly, a hallucination. This option seems to be no more than an *ad hoc* defence of sonicism.

Another option would be to give up sonicism entirely. This is, in my opinion, throwing the baby out with the bathwater. As I admitted earlier in this subsection, sonicism is true when we focus on hearing sounds and sound sources in the causal sense. What we should better do is simply to limit the scope of sonicism to cases of hearing which involve sounds. That said, when it is restricted in this way, in the causal sense of hearing, sonicism is extremely uninformative. As shown by the above example of bone conduction, to say that a non-sound entity is heard *qua* sound source is just to say that it is heard via the causal mediation of its sound. It is then necessary that a sound source *qua* sound source is heard less directly than a sound. Therefore, in the causal sense of hearing, sonicism is an *a priori* truth rather than an empirical discovery.

I am happy to accept this consequence, partly because I am not attracted to sonicism from the very beginning, and partly because limiting it in this way indeed highlights a more interesting and noteworthy fact about hearing: we hear more than sounds and sound sources. It is no longer justified to treat hearing as merely the perception of sound and sound source. I think this improved understanding of hearing is worth the cost of turning sonicism into an uninteresting claim.

We should then focus on cases where a sound is heard. As for hearing in the representational sense, I have already argued in Chapter 2 that it is more plausible that sound sources are direct auditory objects than not. To the extent that the direct auditory object of an auditory experience appears to be distally located, my wave theory does not affect this conclusion and hold that it is the event source.

An auditory experience represents a sound *and* its event source. There does not seem to be any ground to say that either of these auditory objects is represented in virtue of the other being represented. Indeed, both sounds and event sources do not appear to be represented in virtue of any other thing being represented.

Recall that I define the property of representing an *x* as the representational property of being a phenomenally *x*-ish representation. My main reason for saying that event sources are direct auditory objects is that this gives us a simple theory of auditory perception. Focusing on the phenomenal character of an auditory experience, the direct auditory object appears to be distally located at the location which is, in fact, the location of the event source, and its auditory properties covaries systematically with the acoustic properties of the event source. This correspondence supports understanding the auditory experience as a phenomenally event-source-ish representation, and hence as a representation of the event source.

Similarly, in the previous section, I note that the phenomenal character of an auditory experience also covaries systematically with the acoustic properties of the two local parts of the compression wave at the eardrums, and this covariation is different from that with the acoustic properties of the event source. This correspondence supports understanding the auditory experience also as a phenomenally constituent-fluctuations-ish representation, and hence as a representation of the two constituent fluctuations at the eardrums.

These two cases differ from typical cases of indirect representations. When I see the tank is empty by seeing the indicator on the dashboard, the phenomenal character of my visual experience does not covary with properties of the tank but those of the indicator. When I see my supervisor in a video meeting, the phenomenal character of my visual experience does covary with some of his visual properties but only indirectly. The covariation needs to be broken down into two parts: one between the phenomenal character of my experience and the visual properties of the video, and one between the visual properties of the video and the visual properties of my supervisor.

One might object that the covariation between the phenomenal character of an auditory experience and the acoustic properties of the event source should also be broken down into two parts with the local acoustic properties of the compression wave as the bridge. This is phenomenologically inaccurate. An auditory experience represents both intrinsic and perspectival properties of event sources. Earlier I proposed that the perspectival properties are actually local properties of the compression waves, as they covary systematically. There is still another dimension in which the phenomenal character of the auditory experience varies, and this variation matches the variation of the intrinsic acoustic properties of the event source. It is true that our brain needs to work out the intrinsic acoustic properties of the event source from the local acoustic properties of the compression waves. However, this is a subconscious process. At the experiential level, such processing is already over and all we are conscious of is the resulting auditory experience which directly varies with the intrinsic acoustic properties of the event source.

It seems a possible explanation for why people may find sonicism attractive is that they do not, as I do in §2.1.2, distinguish between auditory objects and heard objects. The claim that I would not hear an event source if I do not hear its sound can be understood in different ways. The indisputable one says that I would not have *this* auditory experience which represents an event source if the compression wave produced by that event source did not cause me to have this auditory experience. The problem is that this conflates the causal sense and the representational sense of hearing. Notice that it is also indisputable that I would not have *this* auditory experience which represents a compression wave if that compression wave did not cause me to have this auditory experience. If sonicism does not require the same sense of hearing, then it should also be allowed to say that we hear a sound in virtue of hearing that sound. But I suppose this is not something sonicists would accept. Therefore, when we evaluate sonicism, we should use the same sense of hearing for sounds and event sources. Then my contention is that sonicism is true only in the cause sense of hearing when we focus on sounds and sound sources.

6.7 Against Wave Pattern Theory

In §4.1, I distinguished two major versions of medial theory. The first one is wave theory, which identifies sounds with concrete entities in the medium. My view falls into this category. As for the second one which I called wave pattern theory, it could also be treated as a variant of wave theory. It identifies sounds with patterns or structures of frequency components instantiated by compression waves. In this section, I argue against wave pattern theory.

First, in identifying sounds with abstract entities, wave pattern theory makes the causal role of sounds problematic. Sounds themselves would have no causal power. It is rather the compression waves instantiating them which can produce in us auditory experiences. In contrast, wave theory does not have this consequence. Compression waves and their constituent fluctuations are all concrete events which have the right ontological nature to enter the causal structure of auditory perception.

I can see no compelling reason to identify sounds with abstract patterns or structures of frequency components rather than the compression waves which contain those components. Nudds motivates his wave pattern theory by appealing to the way our auditory system serves its function of perceiving sound sources. He says:

> … the sounds we experience are the result of the way the auditory system groups the frequency components that it detects in order to extract information about the sources that produced them. (Nudds, 2009, p. 75)

This claim would be blatantly objectionable if it means that what sounds objectively are out there in the environment results from the way the auditory system functions. It would be plausible if it instead means that the way we experience sounds to be results from how the auditory system works. However, this would have no direct implication for what sounds are objectively. Moreover, even if this alleged relation between sounds and our experiences of them is granted, it still does not follow that sounds are abstract patterns of frequency components. The very same observation could also be explained if sounds are concrete, patterned combinations of frequency components, in which case sounds can unproblematically be the causes of our auditory experiences. Wave pattern theory is therefore unmotivated.

Second, in Nudds's view, the identity of a sound is fixed by the causal origin of the frequency components instantiating it. Sounds cannot be instances of patterns of frequency components, because people hearing the same sound can be hearing different instances at different locations (ibid., p. 76). Nor can sounds be types of patterns of frequency components, because numerically distinct sounds can be qualitatively identical (ibid.). Different instances of patterns are counted as the same sound only if they are instantiated by frequency components produced by the same event source (Nudds, 2010b, p. 293).

I agree with Nudds regarding the role of sound sources in fixing the identities of sounds. However, my view has a more elegant explanation for the three observations given by Nudds: the identities of compression waves is determined entirely by the identities of their event sources, such that qualitative similarity is simply irrelevant. In my view, constituent fluctuations belong to the same compression wave because of their common causal origin. This is explained by the causal nature of compression waves as events in which pressure fluctuations propagate in the medium. Constituent fluctuations of the same compression waves are not connected by qualitative similarity but by causal chains linking them back to their event sources as the common causes.

As for Nudds's view, it remains unexplained why the identity of a sound, considered as an abstract pattern, is fixed by another property of its instantiator, *viz*. the causal origin of the frequency components instantiating it. While a propagation event necessarily contains a causal link leading back to its event source, there simply does not appear to be any necessary connection between an abstract pattern and the causal origin of its instantiator.

Third and relatedly. If Nudds is right that a sound is just a pattern, it is natural to wonder how dissimilar two constituent fluctuations of the same sound can be. Nudds does not offer any answer to this question, so it is not clear if he thinks there is any limit to this dissimilarity. In contrast, I simply deny the relevance of qualitative similarity to the identities of sounds.

It is implausible to limit how dissimilar a sound could be at two different locations. There are many ways in which the waveform of a compression wave can change dramatically. For example, if the amplitude of a compression wave is high enough, the compressions and the rarefactions are themselves significant changes to the medium which respectively increase and decrease the speed of sound. The compressions and rarefactions of the same wave would then propagate at different speeds, and hence the waveform would change in the process of propagation. Also, the faster rate of attenuation for frequency components at higher frequency means that only the lower frequency ones will remain in the later career of a compression wave. These are examples in which very different patterns of frequency components are instantiated while still being counted as the same compression wave, and hence should be counted as the same sound.

We might even imagine a world in which the medium obeys an entirely different set of physical laws. A compression wave propagating in this medium would change its waveform in whatever imaginable way. Constituent fluctuations at different locations would appear to us like randomly generated, although they are indeed causally connected by the physical laws in that world. Nonetheless, insofar as they are caused by the same event source, they still constitute the same compression wave in my view. However, the patterns of frequency components would be as many as the constituent fluctuations there are.

It is unclear how Nudds would conceive of this imaginary world. If he thought that there is only one sound, then it is an unnecessary complication to identify a sound with infinitely many patterns of frequency components rather than a compression wave constituted by infinitely many patterned local parts. If he instead thought that there are as many sounds as there are patterns of frequency components, then the same judgement should be made regarding sounds in *our* world. This is because the only difference between the two worlds is just the degree of qualitative similarity between constituent fluctuations, and this is accepted by Nudds as neither sufficient nor necessary for the identity of sound. As a result, he would need to give up the plausible idea that people at different locations can hear the same sound.

Overall, wave pattern theory is unmotivated, has inferior explanatory power, and leads to implausible judgement regarding the identity of sound. My view, in contrast, does not have these problems and hence should be preferred.

6.8 Features of Sounds Revisited

In this section, I investigate the implications of my version of wave theory on the eight features of sounds discussed in Chapter 3. Although I dismissed half of them as not theoretically neutral and hence cannot serve as evidence for any theory of sound, I did not deny that they might be true features of sound. However, if any of them turns out to be mistaken, then it would be preferable if a theory contains the necessary resources to give a plausible account of the origin of the error. As we will see in a moment, my theory can accommodate all the eight features, although a few of them are explained not so straightforwardly.

6.8.1 Relation to Auditory Perception

Sounds as compression waves are causal events in which pressure fluctuations propagate in the medium. Recall the distinction between "auditory object" and "heard object" introduced in §2.1.2. A heard object is an external object which causes a perceptual experience in a subject in the right way required for veridical perception. Sounds are heard objects in this sense: we hear them because our auditory experiences are caused by its local parts at my eardrums. We cannot, however, hear the propagation of pressure fluctuations, as we cannot hear the causation between the constituent fluctuations of a compression wave.

As for hearing in the representational sense, sounds are auditory objects in a not so obvious way. In terms of phenomenology, there is no apparent distinction between a sound and its source. It is through reflecting on the experience that we can realise that both the sound and its source are represented.

My view also implies that sounds are not proper objects of hearing. In a rock concert, you can hear *and* feel the bass. Although the phenomenal characters of the auditory experience and the tactile experience are quite different, this does not show that these experiences represent different entities. Nor are sounds the only kind of entities we hear. We also hear sound sources—both event sources and thing sources.

6.8.2 Public Objects

Compression waves are public objects in the sense that different people can hear the same compression wave in virtue of hearing different local parts of it. However, since the local part I hear can only be heard at the location I occupy at the moment of hearing, no two people can hear the same local part at the same time. This does not suggest that local parts are private objects. After all, if someone else *were* here at my location now, she would hear the very same local part I hear. The case is therefore different from typical private objects like pains, for which no such counterfactual would hold.

This suggests a similarity between hearing and touch. When we touch an object, we touch it in virtue of touching the parts which are in contact with our hands. Since hands cannot overlap, different people cannot touch the same part at the same time. Nonetheless, someone else could have touched that part if I were not touching it at that moment. Of course, she can also touch the same part at a different time. This possibility is not available for sounds, but this difference is irrelevant to the issue of publicity. It is simply a difference resulting from the ontology of objects and events: objects and their parts endure through time,[36] while events have different temporal parts at different times.

6.8.3 Temporality

In understanding compression waves as events, I commit myself to the idea that sounds are essentially temporal beings. The temporality of sound has two main aspects: duration and temporal direction. I begin with duration, as it is often the only aspect discussed in the philosophical literature.

In §6.4, I responded to O'Callaghan's challenge that we misrepresent the durations of sounds if wave theory is true. My reply focuses on the content of our auditory experiences. This subsection goes further by determining what temporal properties sounds possess.

Two proponents of wave theory—Kalderon (2018b, pp. 105-106) and O'Shaughnessy (2009, pp. 117-118)—hold that a sound has a "double-duration". One duration roughly matches the duration of the event source. This is the duration

[36] Unless four-dimensionalism is true.

measurable at a fixed location. Another duration is the lifetime of the compression wave. It begins as soon as the event source starts and ends only when the compression wave completely dies out in the medium. We may treat these two durations as the local and global durations respectively.

Wave theory is the only theory which needs to admit that sounds have both global and local durations. For other theories, sounds only have one duration. This fits well with ordinary discourse. Even for a wave theorist like me, if I were being asked about *the* duration of a sound, I would not hesitate to answer what in fact is the local duration. Indeed, in presenting his temporal misrepresentation objection, O'Callaghan (2007b, p. 43) assumes implicitly that sounds have only one duration. The objectionable implication of wave theory, in O'Callaghan's view, is that a sound only has the global duration, but our experience misrepresents it as having the local duration.

Both global duration and local duration require some clarification. The global duration of a compression wave is the duration of the propagation of pressure fluctuation. Since we cannot experience the propagation, we cannot experience the global duration of a sound. Moreover, it is practically impossible to determine the global duration of a sound in ordinary settings. The medium surrounding a sound source is uneven. Such an unevenness affects how far and how fast the pressure fluctuation can propagate along different paths. As a result, a compression wave would not end at the same time in every direction. Considered as a whole, the global duration of a compression wave equals the duration of propagation along the longest path. However, it is practically unpredictable which path is the longest and where it ends, it is therefore impossible to measure the global duration of a compression wave.

In contrast, local duration is what we can measure. In the metaphysical framework presented in §6.5, the local duration of a compression wave is its local feature. It is how long a compression wave happens to the medium at a fixed location. We might also simply say that it is the duration of a constituent fluctuation. Generally, a natural environmental sound does not end abruptly but decay gradually. Since the softer tail cannot propagate as far as the stronger attack, the compression wave has a shorter local duration further away from the sound source. However, if a sound is strong both at the beginning and at the end but has a soft middle part, at a certain distance from the sound source the middle part might have already died out and therefore leave a gap in the middle. We may then say that the compression wave stops happening to the medium at that location during the gap. Shall we then exclude the gap from the local duration of the compression wave there? If so, does it mean that the compression wave splits into two once its middle part dies out?

As I have said in §6.5, compression waves are identified in terms of their event sources. Even though the compression wave stops happening at that location during the gap, the remaining temporal parts are still caused by the same source and hence belong to the same propagation event. The two stronger parts are thus parts of the same constituent fluctuation. The constituent fluctuation is gappy, but it continues even during the gap. Its duration should therefore include the gap.

There is another issue concerning local durations. Reflection of compression waves is practically inevitable. A compression wave does not merely reflect off the walls in a room, but also reflects on the floor and under the ceiling. Indeed, even if we can avoid reflections from walls on an open field, the reflection from the ground is hardly avoidable. This means usually a compression wave arrives at a location around its source along at least two paths. Since the direct path is necessarily shorter than the reflected one, the compression wave begins happening at that location when it arrives through the direct path and ends until the reflection is gone. If we simply identify the local duration of the compression wave at that location as the total duration in which the compression wave happens there, the local duration is longer than the duration of the event source. The difference may be negligible on an open field, but it could be significant in highly reverberant spaces such as cathedrals.

Considered in itself, this consequence might not be objectionable. After all, if we allow the local duration of a compression wave to be shorter than the duration of its event source at a long distance, then why can't it be longer in some cases? However, a problem arises if we take into account the earlier conclusion that a constituent fluctuation can be gappy. Let us consider an echo which arrives at a location after the primary sound ends. The echo and the primary sound are indeed the same compression wave. The situation is fundamentally the same as the case of ground reflection: the compression wave arrives at that location via two different paths. This time, the direct path is shorter than the reflected one again but significantly. As a result, the direct sound and the reflection does not overlap temporally. However, since they are caused by the same event source, they are parts of the same propagation event. It seems then I should be consistent and say that the constituent fluctuation at that location is gappy and has a duration much longer than that of the event source.

I do not find this conclusion satisfactory. It seems there is a sense in which the primary sound and the echo should be treated as distinct individuals, but the notion of constituent fluctuation in the current sense cannot capture it. Originally, the notion of constituent fluctuation is introduced in a context where reflective surfaces are not in consideration. In other words, every constituent fluctuation in that context is an effect of a compression wave arriving at their location through a direct path. It is intended to be a part of a compression wave which is determined solely by its event source and the medium. The properties of a constituent fluctuation are then correlated to the properties of its event source in a relatively simple way. How does the presence of reflective surfaces disrupt this correlation?

The problem is not that reflective surfaces are not perfectly reflective. It is true that actual reflective surfaces affect the properties of the reflected waves, but even if we consider the idealised situation where all reflective surfaces are perfectly reflective and thus leave no trace on the reflected waves, the problem remains.

The problem seems to be rather that the presence of reflective surfaces makes it possible for an event source to affect the pressure fluctuation at a location via multiple paths. In a confined room, every location in it is connected to the event source through a direct path. The effect of the event source on the pressure fluctuation there via the direct path should be the same as the idealised case where no reflective surface is present, and hence should not be affected by the reflected wave from other paths. This effect is what the notion of constituent fluctuation originally intended to capture. We might, therefore, refine the notion accordingly: a constituent fluctuation at a location is the effect of an event source at that location via one causal path. The causal path can be direct or indirect, so this formulation makes room for us to identify constituent fluctuations caused by their event sources via indirect paths involving reflection.

Although this refinement of the notion of constituent fluctuation looks quite reasonable, I shall reject it for two reasons. First, it cannot explain the following case of echo. Normally, a causal path entering an ear (or a measurement tool such as a microphone) does not turn around.[37] Therefore, even when the thing source, the hearer, and the reflective surface is perfectly aligned on a straight line, the reflected wave is guaranteed to be coming along a path different from that of the direct path thanks to the diffraction of the compression wave around the hearer in the first encounter. However, in the absence of any measurement equipment, a compression wave can pass through the same location multiple times along the same path in virtue of reflections. The restriction to a single causal path thus cannot differentiate the primary sound and the echo(es) in this situation.

The second problem concerns the source-motion Doppler effect. When a thing source moves, the part of the medium in contact with it changes constantly. This means the causal path connecting the hearer to the event source at an earlier moment starts at a different point from the one at a later moment. Therefore, the pressure fluctuation detectable at the location of the hearer is caused by the event source via a different casual path at every moment. If a constituent fluctuation is tied to one particular causal path, this means there is one constituent fluctuation for each causal path. Since the movement of the thing source in the medium is continuous, it follows that the pressure fluctuation at the hearer's location consists of an infinite number of instantaneous constituent fluctuations. However, it is implausible that the number of constituent fluctuation at a location would change from one to infinity just because the thing source starts moving relative to the medium.

I shall refine the notion of constituent fluctuation in another way. Let us focus again on the case of reflection. Although the difference between causal paths cannot directly explain the difference between two constituent fluctuations produced by the same event source, it is nonetheless related to another important difference. Since different causal paths have different lengths, at any fixed location the constituent fluctuations produced by the same event source via different causal paths begin at different times. In other words, at any moment those constituent fluctuations are produced by different temporal parts of the same event source. Accordingly, a constituent fluctuation cannot be produced by two temporal parts of its event source simultaneously.

---

[37] More accurately, although a compression wave can reflect from the eardrum and resonate at a certain frequency determined by the length of the ear canal, the reflected wave which manages to escape the ear is too weak to reach the reflective surfaces in the environment.

Since the new refinement has no limitation on the causal path through which the event source produces the constituent fluctuation at a location, the previous problem concerning the source-motion Doppler effect is resolved. However, a new problem arises. If an event source happens in a circular tunnel, the compression wave propagates toward opposite directions. There is a point on the opposite side of the tunnel where the pressure fluctuation is caused by the same temporal part of the event source via two different causal paths. It seems to be reasonable to say that the compression wave overlaps with itself at that location, but this would require there to be more than one constituent fluctuation.[38] However, if the different causal paths do not matter to the number of constituent fluctuation, and the pressure fluctuation at that location is produced by the same temporal part of the event, it seems there is no basis on which we can identify two constituent fluctuations there.

To solve this problem, we may compare this case with another scenario. Compression waves diffract significantly. When there is an object on the path of its propagation, a compression wave just goes around it. If the object is small relative to the wavelength, the object almost has no effect on the propagation of the wave. In contrast, if the object is larger than the wavelength, then it can cast an acoustic shadow behind it, where the amplitude of the compression wave is weaker. It seems the circular tunnel scenario is structurally similar to the more ordinary case of diffraction, where a large object is on a compression wave's path of propagation. If so, then if there should be two constituent fluctuations at the point directly opposite to the event source in the circular tunnel, there should also be two constituent fluctuations at the back of the large object in the ordinary case of diffraction. However, this is not the case. There should be some difference between these two cases which can explain our different judgement.

The crucial difference seems to be that in the ordinary case of diffraction, we do not think that the compression wave overlaps with itself. Without the need to understand the situation in terms of the superposition of different parts of the same compression wave, there is, therefore, no need to distinguish two constituent fluctuations. If we consider the case of the circular tunnel again, we can see that apart from the point equidistant from the event source via the two causal paths, other locations are connected to the event source via causal paths of different

[38] The superposition of compression waves will be discussed in §6.9.2.

lengths. As a result, the pressure fluctuations at those locations are caused by different temporal parts of the event source, and hence there are more than one constituent fluctuations. Therefore, to determine how many constituent fluctuations there are at a location, it is sometimes needed to consider the nearby locations. This makes it very difficult to precisely define the identity condition of constituent fluctuation. Nonetheless, the second refined notion of constituent fluctuation proposed above should work well enough in most situations, and I will employ it in the remaining parts.

According to the refined sense, in the case of an open field, there are two different but temporally overlapping constituent fluctuations at a location. The compression wave, therefore, has two local durations there which are identical to the durations of the two constituent fluctuations. We do not experience them separately because the reflected wave is suppressed as an effect of echo suppression. This allows us to better perceive the event source, as we would not experience two sources while there is only one. Nonetheless, the ground reflection is not entirely discarded in auditory processing. Rather, it makes a difference to the perceived qualities of the event source and hence the constituent fluctuation from the direct path. In this minimal sense, the constituent fluctuation from the indirect path is also represented, though it might not be appropriate to say that it is experienced.

The qualitative differences resulting from reflected waves is much more significant in reverberant spaces like cathedrals because there are much more indirect paths connecting a location to an event source. As a result, there are not only two but many constituent fluctuations overlapping temporally at a location. We may even be able to notice that the auditorily perceived duration of the event source is longer than its actual duration because of a mismatch with our visual experience of it. In terms of metaphysics, the perceived duration is not the local duration of the compression wave, because it is the total duration of all the constituent fluctuations combined. This brings out the fact that although the total fluctuations which cause our auditory experiences are composed of constituent fluctuations, this does not guarantee that constituent fluctuations are experienced separately.

Our auditory system works in accordance with certain principles in analysing the detected total fluctuation into constituent fluctuations, but this process is not perfect. Indeed, it is shaped by the selective pressure imposed by sound sources rather than constituent fluctuations themselves. To the extent that it can help us identify sound sources in the environment well enough, there is no further pressure for it to separate every constituent fluctuation in the total fluctuations detected. Nonetheless, as long as the detected total fluctuations can be differentiated into components roughly matching the constituent fluctuations of the same compression waves, we should not take this imperfection as showing that our auditory experience represents total fluctuations rather than constituent fluctuations. Constituent fluctuations may be misrepresented, but misrepresentations are still representations.

As for the case of distinctive echo experiences, the echo is experienced. That is, both the constituent fluctuation from the direct path and that from an indirect path are experienced. The echo is experienced separately from the primary sound. This means the respective constituent fluctuations are experienced separately. The two resulting experiences both represent the common event source, though the echo experience misrepresents its spatial and temporal location. The auditory properties of the event source may also be treated as distorted in the echo experience because of the imperfect reflection of compression waves. However, the experienced auditory properties in the echo experience *are* local features of the compression wave. The veridicality of echo experiences differs with respect to whether they are considered as representations of event sources or compression waves.

Notice that a compression wave can have very different local durations. In the case of the source-motion Doppler effect, the local durations of a compression wave are shorter in the front of the sound source and longer at the back. Depending on the speed of the sound source relative to the medium, the difference can be enormous. This also affects the rate at which the auditory properties of a constituent fluctuation change: the shorter the duration, the faster the rate of change.

The source-motion Doppler effect also has an important implication on the temporal direction of sound. It might be an attractive idea that the identity of a sound is tied to how its properties change over time (O'Callaghan, 2007b, pp. 22-23). This idea, however, would be wrong in my view if the change must be in one specific direction. Consider the extreme case of source-motion Doppler effect where the thing source is supersonic. This has an interesting result regarding the temporal direction of sound: in the front of the thing source, compression waves arrive at an order opposite to the order of generation. This is because whenever a supersonic thing source compresses the medium, it is at a location not yet reached by the previous compression, and hence later compressions always arrive earlier than their predecessors at locations in front of the thing source.

The reversal in order, however, does not happen at the back of the thing source: the Doppler effect only results in a lengthening of the wavelength. This means the temporal order of the same sequence of compression waves is different at different locations relative to the thing source. Since the same compression wave is the same sound in my view, this means the identity of a sound is not altered by the temporal direction of the changes in its properties. I have argued in §6.5 and §6.7 that the identity of a sound is not fixed qualitatively. However, even if the identity of a sound is tied to the change in its properties, the change should be allowed to happen in either direction.

In sum, my view says that sounds as compression waves are essentially temporal beings. It has so-called "double-duration". The global duration is the duration of the propagation of pressure fluctuation throughout the affected region. In contrast, the local duration is the duration in which the pressure fluctuation at a location is affected by the event source. This means a compression wave has an infinite number of local durations, each at a location in the continuous volume of the medium. Besides, the properties of a sound can change in any temporal direction and at any rate without altering the identity of the sound.

6.8.4 Spatiality

The nature of compression waves as propagations of pressure fluctuation implies that sounds are no less spatial than temporal: propagations can happen only across a region of space. Regarding the spatiality of sound, O'Shaughnessy (2009, p. 118) says that "a sound does not possess a double-location as it does a double-time". The region occupied by a compression wave expands during the propagation. This expanding region is the only location of the compression wave. The spatiality and temporality of sound are therefore disanalogous. Is he correct?

Our discussion on the double-duration of sound helps us to see that there is no such a disanalogy. The crucial idea is that we should distinguish the global location of the entire propagation event from the instantaneous locations of its affected region. A compression wave as a propagation event has a global location encompassing every location the pressure fluctuation of which is affected by the event source, but at each moment there is only a smaller region occupied by its constituent fluctuations. Similarly, a compression wave has a global duration which is the time required for it to happen progressively throughout its global location. At each location within that region, however, there is a shorter period during which the pressure fluctuation there is affected by the event source. If we accept that compression waves have double-duration, then we should by parity accept that they have double-location.

Our auditory experiences represent constituent fluctuations but not the entire propagation events. I argued in §3.4 that auditory objects are either represented as located at or near sound sources or not spatially represented at all. The former option would imply that we misrepresent the locations of constituent fluctuations, since we can only hear those at our eardrums. My view avoids this consequence in virtue of denying that the locations of constituent fluctuations are represented in auditory experiences. Hearing a constituent fluctuation furnishes us with information about the location of its event source, but its own location is not experienced. It is only through reflection that its location is revealed to us.

6.8.5 Relation to Ordinary Material Objects

The idea we examined in §3.5 is that sounds are distinct and independent from ordinary material objects. It is obvious that this is true in my view. Compression waves, once generated, are distinct from their thing sources, and they do not depend on their sources for their persistence.

The conclusion in §3.5, however, concerns not the objective fact but how sounds are represented in our auditory experiences. It says that sounds are not represented as distinct and independent from ordinary material objects. In terms of phenomenology, it seems we only experience one entity which is distally located, and this is the event source. The fact that compression waves are also represented in auditory experiences is a fact revealed only through reflection.

6.8.6 Being Caused by Sound Sources

Compression waves as propagation events are individuated by their event sources. An event source causes pressure fluctuations in its immediate surrounding. Such pressure fluctuations then cause pressure fluctuations at locations further away from the sound source. All these local pressure fluctuations constitute a propagation event in virtue of their causal connection. This propagation event is thus caused by the event source in the sense of being initiated by it.

Individuating compression waves by reference to event sources allows a certain degree of vagueness. Consider this example. Using the right hand to pluck a string on an acoustic guitar produces a sound. Using the left hand to press on the same string and slide across the frets also produces a sound (but a weak one). How many sounds are there if I perform these two actions at the same time? Normally we would say that there is a sound rising in pitch. We may call it the sound of plucking and sliding simultaneously. However, if the slide stops before the string stops vibrating, it seems we can distinguish the sound of plucking from the sound of sliding—the latter sound is shorter and ends earlier than the former one. Therefore, both one sound and two sounds are reasonable answers.

My view can accommodate this vagueness in the number of sounds by appealing to the vagueness in the number of event sources. To the extent that it makes sense to say that there are two event sources—plucking *and* sliding, it makes sense to say that there are two sounds. Similarly, to the extent that it makes sense to say that there is only one event source—plucking *while* sliding, it also makes sense to say that there is one sound only.

The status of constituent fluctuations as causal effects of event sources provides a veridical condition for experiences of compression waves. In normal situations, there are multiple sound sources in the environment, such that at each location multiple compression waves overlap with each other. The total fluctuation at a location thus contains frequency components from all the overlapping compression waves present there. Considered in themselves, these frequency components have no special relation to each other. Therefore, if our auditory experiences represent total fluctuations, insofar as all the frequency components present in a total fluctuation is represented, it is equally well to represent it as a single individual or represent each frequency component separately.

However, our auditory experiences represent constituent fluctuations. Constituent fluctuations, as we have seen repeatedly, are constituents of compression waves. If multiple compression waves overlap at a location, there are multiple constituent fluctuations. The total fluctuation at a location is the sum of all constituent fluctuations there. There is an objective fact about how many constituent fluctuations there are at a location.

Our auditory experiences can represent constituent fluctuations veridically. Frequency components of the detected total fluctuation are analysed by our auditory system into separate groups. In ideal situations, such groups correspond to each of the constituent fluctuations there, and hence the constituent fluctuations are represented veridically. How frequency components are grouped is shaped by the function of hearing to perceive event sources. Though this does not afford perfect differentiation of total fluctuations into constituent fluctuations, it is nonetheless true to say that sound source perception explains how our auditory experiences manage to represent constituent fluctuations.

6.8.7 As Bearers of Auditory Properties

Initially, I claimed in §6.1.2 that sound sources inherit their auditory properties from the sounds produced. Sounds are then the primary bearers of auditory properties. Now, after presenting my version of wave theory, some adjustment to this claim is needed.

To begin with, we shall consider the claim that sounds bear auditory properties. I distinguished the global, local, and instantaneous features of compression waves in §6.5. Which of these kinds should auditory properties belong to? Since auditory properties are what our auditory experiences represent sounds as having, and we hear sounds by hearing their local parts, auditory properties should, therefore, be local features of sounds. Since a compression wave has different local features at different locations, it is then unclear how sound sources inherit auditory properties from their sounds.

A possible suggestion is that we may attribute global auditory properties to compression waves based on their local auditory properties. Pitch, it might be suggested, could be attributed in this way. As the frequency of a sound would not change with distance, so there is a pitch common to every constituent fluctuation of a compression wave, and hence we can say that the compression wave as a whole has the same pitch as its constituent fluctuations.

This method does not apply to loudness and timbre, as these two properties obviously vary across locations. Even in the case of pitch, the above suggestion mistakenly assumes that the invariance of frequency with distance implies a pitch common to every constituent fluctuation. Consider the case of the source-motion Doppler effect. At different angles from a sound source moving relative to the medium, the constituent fluctuations of the same compression wave have different frequencies. There is thus not a single pitch shared by all constituent fluctuations which could be attributed to the entire compression wave.

It might be replied that the source-motion Doppler effect distorts our perception and hence should be avoided in judging the pitch of compression waves. Therefore, the source-motion Doppler effect does not show that we cannot employ the suggested method in attributing global pitch to a compression wave if the thing source is stationary. This response does not work. First, notice that not only are there many moving sound sources, the presence of wind is also a cause of relative motion between sound sources and the medium. This makes the suggested method inapplicable in most situations except a very limited range of highly artificially controlled cases. Moreover, while the source-motion Doppler effect does distort our perception of the event source, our perception of the compression wave is not distorted. Nor is the compression wave itself in any sense distorted. Although a compression wave produced by a moving thing source differs from what it would be if the thing source is stationary, it is no less a compression wave. Considered as such, there is no normative standard governing how compression waves should be, and hence there is no reason to treat cases differently depending on whether the thing source is stationary or not.

We may take a further step to argue for the stronger claim that the suggested method of attributing global pitch to compression waves is unjustifiable in principle. To see the reason, let us consider the case of gravitational redshift. A clock ticking on a more massive body is slower for an observer on a less massive body. Similarly, for a light source on a more massive body, the frequency of the photons emitted is lower if the measurement is done on a less massive body. Since a lower frequency means a longer wavelength, the light as observed on that less massive body is redshifted. For ease of discussion, we may exaggerate the effect and say that the light is blue for the observers on the more massive body but red for the observers on the less massive body. We should therefore relativise the colour of light to the location of observation.

Back to the case of pitch. Pitch is mainly correlated to frequency. We can imagine a possible world in which compression waves can propagate as far as electromagnetic waves do. In this world, the same reason for relativising the colour of light to the location of observation would apply to the case of compression waves.

Since this acoustic counterpart effect of gravitational redshift exists also when sound sources do not move relative to the medium, it provides a more general reason for not treating pitch as a global property of compression waves.

We may nonetheless compare two compression waves in terms of auditory properties. If the conditions for observation are strictly controlled, it is possible to say that a compression wave is, say, louder than another one. But the real meaning of this comparative statement is that the *local* loudness of the former compression wave is louder than that of the latter one *at every location*. Therefore, although such a global comparison between compression waves is possible, it does not follow that auditory properties can be global properties of compression waves.

I admit that it is a counterintuitive consequence of my view that sounds as compression waves do not have global auditory properties. It is typical to attribute auditory properties to sounds without relativising to any location. However, I do not think that my view is therefore implausible, because the relativisation of auditory properties to location is actually quite common. In an orchestra, the sound of the trumpets is deafening for the bassoonists sitting in front of the trumpet section, but the trumpetists would not hear the sound of their instruments as loud as the bassoonists do. Is the sound of the trumpets loud? It is loud at the bassoonists' location but not at the trumpetists' location.

It is worth noting further that the relativised judgements are more often made when we are in the more cautious moments. Enforcement of noise control ordinance, for example, requires that the noise level should be measured at a specific distance from the sound source. Therefore, it seems even ordinary practice does not conflict with relativising auditory properties to location.

I shall close this subsection with a note on attributing auditory properties to sound sources. Unlike the case of compression waves, I think it is reasonable to privilege the constituent fluctuations closest to sound sources and attribute auditory properties to sound sources with reference to them. While every constituent fluctuation of the same compression wave is equal in being its constituent, some constituent fluctuations preserve information about their source better. Wave propagation is a process of information transmission, and it is not lossless. The amount of information preserved should, therefore, serve as a normative standard for deciding which constituent fluctuations perform the role of perceptual means the best.

The source-motion Doppler effect, however, implies that even those constituent fluctuations immediately around a sound source can differ with each other. Unlike the case of compression waves where there is no standard governing how they and their constituent fluctuations should be, it is fine to say that some perceptual conditions should be preferred for the attribution of auditory properties to sound sources. It makes perfect sense to say that the auditory properties of a sound source should match the auditory properties of its sound in ideal observation condition, in which, at least, the Doppler effect should be absent.

Privileging the constituent fluctuations immediate around their event sources implies that we normally do not hear event sources via the most faithful means. However, this need not be a problem for my view. Perception faces more constraints than fidelity. It is impractical to hear everything by placing our ears on them. Some sound sources are too loud to be heard closely. Some sound sources are too dangerous: do not hear the roar of a tiger at its throat. Our auditory system can extract from less faithful constituent fluctuations enough information for our survival. To the extent that my view does not render the evolutionary fitness conferred by our hearing an unexplainable myth, the metaphysical question about sound should be kept separate from the problem of perception.

6.8.8 Survivalism

Survivalism is the claim that sounds survive qualitative changes. This could be understood in two ways if sounds are compression waves. First, it focuses on the local level and means that each constituent fluctuation survives changes to their properties. Second, it focuses on the global level and means that a compression wave survives changes of its local features through its propagation.

There is a simple reason to accept survivalism in both senses in my view: compression waves are not identified qualitatively. A compression wave is just a causal event in which local pressure fluctuations are propagated into the medium. At the global level, constituent fluctuations belong to the same compression wave not because of their qualitative similarity but because of their common causal origin. Therefore, qualitative difference between constituent fluctuations is irrelevant to the identities of compression waves.

Similarly, at the local level, a constituent fluctuation at a location is the effect of its event source there. No matter how its properties change across time, it maintains its identity insofar as the qualitatively different temporal parts are caused by the same event source.

6.9 Special Cases

With a few exceptions, our discussion so far mainly concerns sounds and sound sources without considering the interaction of compression waves with the environment. In this last section, I explore how my view can explain acoustic phenomena beyond such simplest scenarios. First, in §6.9.1, I discuss the reflection of compression waves in the presence of reflective surfaces. Next, §6.9.2 examines the interference of compression waves in the presence of more than one sound source. These two phenomena are then considered together in §6.9.3 to explain the case of resonance in a wind instrument such as the oboe. Finally, in §6.9.4, I move on to discuss an artificial phenomenon—the reproduction of sounds using audio recordings.

6.9.1 Reflection: Echo, Reverberation, Echolocation

Echo has been discussed in relation to our echo experiences a few times in previous chapters. The main idea is that locating sounds at or near their sources and treating echo experiences as distorted or illusory experiences of primary sounds fails to do justice to the objective content of our echo experiences. This shortcoming of distal theory can be avoided by wave theory.

Wave theory, in general, allows that our echo experiences are veridical experiences of the reflected waves, as the common agreement that sounds are auditory objects just means that compression waves are auditory objects. I can hear the primary wave and the reflected wave, both of which are correctly experienced when they arrive at my location. Indeed, the two experiences are experiences of the same compression wave at different stages. For those who accept also that sound sources are experienced auditorily, both my experiences of the primary wave and that of the reflected wave represent the sound source as well.

The echo experience might be illusory insofar as it may mislead me about the temporal and spatial location of the sound source. However, echo experiences, at least the distinctive ones which represent the echoes as causally related to the primary sounds, have their distinctive phenomenal characters, such that normal adults would not be misled by them. The case is like our visual experiences of mirror images. Such experiences can be misleading, but only occasionally would we be deceived. As soon as we look around, the different ways in which the worlds “inside” and “outside” the mirror change their appearances in response to our movement immediately lead us to correctly interpret our experiences.

My view is not an ordinary wave theory. I hold that our auditory experiences do not represent the entire compression wave but its constituent fluctuations. An experience of the primary sound thus represents the first constituent fluctuation at my location which is caused by the event source. An echo experience then represents another constituent fluctuation at my location which is also caused by that event source. If this is a distinctive echo experience, it further represents this constituent fluctuation as causally related to the constituent fluctuation represented in the corresponding experience of primary sound.

It is not clear how this causal relation is represented. The later constituent fluctuation may be represented as caused by the earlier one, or simply as being caused by the same source. I cannot determine which option better characterises the phenomenology of distinctive echo experiences. Anyway, both options allow distinctive echo experiences to be veridical, as constituent fluctuations of the same compression wave do share the same causal origin, and the later ones are caused by earlier ones.

Notice that not every echo leads to echo experiences. As I mentioned in §4.2.1, there is an effect called “echo suppression” which happens when the delay between the arrival of a primary sound and the arrival of its echo is shorter than a certain threshold. When it happens, our auditory system fuses the echo with the primary sound, such that the resulting experience represents the primary sound (and the event source) but with slightly different qualities. The qualitative difference between hearing your singing in the bathroom and on an open field is a result of this process. Ordinarily, we talk about such a difference in terms of the reverberation of a venue. To the extent that we can distinguish the reverberation of a venue from the event source, our experiences of primary sound still represent the echoes but only obscurely.

One question arises here: do we auditorily experience the venue? If we can hear the difference between singing in the bathroom and singing on an open field, it seems implausible to say that the venue is entirely absent in our experience. However, Young (2017) takes our experiences of reverberation as showing that we can hear spaces, but he also holds that they fall short of representing the venues. It is unclear how theories other than wave theory would consider this possibility.

My view need not and cannot give a determinate judgement on this issue, but it should be compatible with the possibility that both venues and spaces are also represented in our auditory experiences. At least, a compression wave in my view can carry information about the venue and space in the surrounding in virtue of being modified by the enclosing surfaces. However, it is not explicit in our auditory experiences that venues and spaces are represented alongside event sources. We may nonetheless learn through reflection that some of the qualities of the experiences can only be explained by the contribution from the venues and the spaces. Therefore, the case is like how we learn that constituent fluctuations are represented in our auditory experiences, though it seems it is even more difficult to distinguish features of venues and spaces from features of event sources.

We should not generalise the case of echo and reverberation to all objects which reflect compression waves. We discussed human echolocation in §3.3 and §4.2.4.2. The impressive performance of human echolocators in locating reflective objects strongly suggests that such objects are represented in their experiences. Indeed, following the instructions of Schwitzgebel and Gordon (2011, p. 61), I tried to echolocate my hand and can attest to the auditory phenomenal difference between hearing my silent hand held at different positions. I did not merely hear the clicks I made with my tongue as being modified by my hand in different ways. There was also a clear impression of some silent object located fairly determinately at a certain nearby location, a location which is, in fact, the location of my hand.

I do not, however, pretend to have settled the question concerning whether reflective objects are represented in auditory experiences via echolocation. Echolocatory experiences are far from being familiar to most people. It would not be surprising to find out that I am somehow abnormal in this respect. I should, therefore, end this subsection merely by a remark on how my theory of sound could explain echolocation.

Compression waves are, in my view, propagation events in which constituent fluctuations are causally connected. This claim leaves room for objects in the environment to be involved in the causal process. Right next to a reflective wall, after the compression wave is reflected, the constituent fluctuation there is both caused by its predecessor at the same place and the wall's response. As a result, the constituent fluctuation after reflection contains information about both the event source and the wall.

In general, reflective objects do not all reflect compression waves in the same way. They absorb energy. The smoothness of their surfaces in comparison to the wavelengths of the incoming wave affect the reflection in a way similar to how light is reflected by matte or glossy surfaces. Since the smoothness of a surface is relative to wavelength, this means the same surface can be smooth for lower frequency but rough for higher frequency. As a result, the spectrum of the constituent fluctuation after reflection is altered by the reflective wall, and therefore the characters of reflective objects are best revealed by broad-spectrum noise. This is similar to the case of light, where reflective objects have different reflectance properties. For the same reason as the colours of objects are determined under full-spectrum illumination (e.g. sunlight), blind people also make use of broad-spectrum noises (e.g. white noise, clicks) for echolocation.

There is, however, a dissimilarity between visual and auditory experiences of reflective objects. Very often, light sources and reflective objects are not within the same view. Our visual system, therefore, needs to extract information about the character of light sources from the invariances among multiple reflective objects in the same view. In contrast, echolocators always hear the sound sources as well. Therefore, at least at the level of subconscious auditory processes, it is plausible that the primary sound and the reflected sound can be compared. It would not be surprising if reflective objects are processed somewhere along the auditory pathway and figure in some way in our auditory experiences. If this is the case, my view would then need to say that our auditory experiences represent reflective objects in virtue of representing two constituent fluctuations, although phenomenologically speaking they are not represented as two distinct individuals.

It is still an open question whether the phenomenology of an echolocatory experience is more like a modified experience of the event source or an experience of two explicitly distinct individuals, i.e. the event source and the reflective object. Both options are compatible with my view, as they are consistent with the fact that the stimulus is a local part of a modified compression wave, the modification of which could be identified by comparing the modified and the original features of the same compression wave.

6.9.2 Interference: Constructive, Destructive, Beating

When two compression waves meet, they pass through each other. The amplitude of the total pressure fluctuation at each location within the overlapping area is the sum of the amplitudes of the constituent fluctuations of the overlapping waves. Constructive interference happens when two compressions or two rarefactions add up. If a compression instead overlaps with a rarefaction, then destructive interference happens.

The effect of constructive and destructive interference is reflected in the perceived loudness of sounds. The case of constructive interference is less interesting, as an increase in loudness is what we would intuitively expect when there are more than one sounds. I should, therefore, focus on the more perplexing case of destructive interference—more specifically the case in which two sounds completely cancel each other.

It may be suggested that since there is no pressure fluctuation at the point where two compression waves completely cancel each other, wave theory would imply that there is no sound. Accordingly, a perceiver at that location hears no sound because there is none to be heard. I do not think that wave theory—at least my version—would have such an implication. Using my terminologies, sounds are identified with compression waves rather than *total* fluctuations. The presence of multiple compression waves in an area implies that there are multiple *constituent* fluctuations at each location within that area. Indeed, in the case of complete destructive interference, it is the presence of multiple constituent fluctuations which explains the absence of total fluctuation. Although our auditory experiences represent constituent fluctuations rather than total fluctuations, the absence of total fluctuation prevents our auditory system from detecting the constituent fluctuations there and hence we fail to experience them. It is precisely because there are multiple sounds at a location that none of them can be heard. If the perceiver walks around, she could hear the sounds at some other locations. Therefore, interference—destructive and constructive—only affects the audibility of sounds.

Distal event theory offers a similar diagnosis of the case. There are multiple sounds, each of which is located at the respective sound source. Compression waves are the means by which we hear distally located sounds. The behaviours of compression waves, including how they interfere with each other, do not affect the sounds. Rather, they merely distort our perception and hence affect the audibility of sounds. Therefore, distal event theory would judge that there are multiple sounds in the case of complete destructive interference, and they are rendered inaudible by the distortive behaviour of compression waves as perceptual means.

It seems the explanations offered by wave theory and distal event theory are equally well. This should also be the case for other interference phenomena, such as beating. Typically, when people discuss complete destructive interference, the example concerns two sine waves which have the same frequency and amplitude but are 180° out of phase. But interference happens also in other situations. If there are two sine waves which have different but similar frequencies at a location, such as 439 Hz and 441 Hz, they interfere in such a way that the frequency of the combined wave at that location is the average of the two sine waves (i.e. 440 Hz), and the local amplitude of this wave fluctuates at a rate equal to the difference between the frequencies of the two sine waves (i.e. 2 Hz). The two sounds would not be heard apart but as one sound which rises and falls in loudness.[39]

It might be thought that wave theory would say that there is a compression wave at 440 Hz which fluctuates in amplitude. This is false in my view. At each location of the overlapping area of the two sine waves, there is a total fluctuation containing two constituent fluctuations with constant amplitude, one at 439 Hz and one at 441 Hz. The total frequency, on the other hand, is at 440 Hz and has a fluctuating amplitude. Since compression waves are constituted not by total fluctuations but by constituent fluctuations, there is no compression wave the local properties of which match the properties of the total fluctuations. In other words, there are only two sounds which are the two sine waves.

The auditory experience, in this case, is caused by the two total fluctuations at the eardrums. However, our auditory system is unable to analyse the detected total fluctuations into constituent fluctuations. Instead, it mistakes each of the two total fluctuations as composed of only one constituent fluctuation the frequency and amplitude of which equal those of the total fluctuation. As a result, the experience misrepresents the two sounds as one.

Likewise, distal event theory would also say that there are two sounds at different frequencies and constant amplitudes, and they are located at their sources.

---

[39] To simplify the case, I assume that the two sound sources are located close together such that they cannot be heard apart in virtue of spatial cues.

Since compression waves are causal products of sounds in this view, their interference does not affect the properties of sounds but our perception of them, resulting in an experience which misrepresents them as one sound.

Both theories agree that there are only two sounds and that they are misrepresented in our experience. Distal event theory only says that our perception of the two sounds are distorted. It remains neutral as to whether the distortion is explained by the compression waves or our perceptual ability. My view, in contrast, specifically points to the inability of our auditory system to analyse the detected total fluctuation in explaining the misrepresentation. This is plausible because if the difference frequency between the two sine waves is significantly larger, we can hear the two sounds apart, but there is no fundamental difference with the beating case at the physical level. Therefore, the misrepresentation should be explained by our auditory system. To the extent that my view offers a more specific and plausible explanation, it is preferable.

6.9.3 Resonance: The Oboe

The oboe is an example I used in §5.2.3 to challenge disturbance event theory. The problem is that disturbance event theory cannot satisfactorily answer the question concerning where the sound as a disturbance event is located. The identification of sounds with disturbance events requires sounds to be independent of the transmission events of compression waves. However, this does not allow the sound of the oboe to be located at the reed, as the vibration of the reed is partly determined by the compression wave reflected from the bell. Neither can the sound be located at the bell, because sounds could be picked up by a microphone inserted inside the bore, where the vibrating air column should be counted as a part of the sound source rather than a disturbed medium in this case.

For my view, there is no difficulty in answering this question. My view allows the distinction between disturbance events and transmission events to be drawn in such a way that the former need not be independent of the latter. The air column vibrates because there is a standing wave formed by the superposition of the compression wave generated at the reed and its reflections. Since I identify sounds with compression waves, the presence of a standing wave implies that there is a sound. This is why sounds could be picked up by a mic in the bore. Of course, the compression wave also escapes from the instrument and fills the concert hall. It is the sound heard by the audience. So, in my view, the sound is located in the air both inside and outside the bore.

There is, however, a problem for my view following my analysis of the case of reflection and interference. To better understand the problem, let us consider how the oboe works again, but with special focus on reflection and interference this time.

To begin with, we need to have a basic idea of what happens at the interface between two media with different impedances. When a compression wave reaches such an interface, part of it is transmitted through the interface while the remaining part is reflected. It might be tempting to describe the compression wave as splitting into two in this situation, but according to my view, the transmitted part and the reflected part are still parts of the same compression wave, as they are produced by the same event source.

The oboe is a conical tube with an open end at the bell. At the open end, there is an impedance mismatch between the air inside and outside the bore, and hence there is an interface where compression waves reflect or pass through. The reflected part of a compression wave meets the upcoming part and propagates back to the reed. It reflects once again at the reed, and this process keeps going on until its energy runs out. The pressure fluctuation inside the bore is determined by the superposition of all the parts of the compression waves in it.

An interesting feature of the case of the oboe is that the pressure difference between the mouth of the player and the bore determines when the reed opens and thereby sets the frequencies of the vibrations of the reed and the air column inside. In my terminologies, the pressure difference is partly determined by the total fluctuation at the reed. This means the interference of compression waves plays a role in the process. In the playing condition, the pressure inside the player's mouth is higher than that at the reed. A combined compression at the reed reduces the pressure difference and opens the reed, and *vice versa* for a combined rarefaction. Moreover, the amplitude of the combined compression needs to reach a certain level to open the reed. This is then related to whether the constituent fluctuations at the reed interfere constructively or destructively. The ratio between the length of the core and the wavelengths of the compression waves then determines the resonant frequencies at which the constituent fluctuations at the bore can interfere constructively. This is why the pitch range of a wind instrument is determined by the length of its bore.

If a compression wave can reflect several times in the bore, the superposition of multiple parts of the same wave can increase the amplitude significantly. This explains why resonance can amplify a sound. However, to let the audience hear the music, the instrument also needs to let its sounds out, i.e. to radiate the compression wave at the bell. This leads to a problem: the more sound is radiated, the less sound is reflected in the bore and hence the resonance is weaker. Moreover, for a given diameter of the tube, the ratio between the radiated sound and the reflected sound is frequency-dependent: more sound is radiated at higher frequency. This means the reflected, unreflected, and transmitted parts of the wave have different spectrums. When we take into account factors such as the attenuation due to the distance of propagation, the damping caused by the friction with the wall of the instrument and the heat loss to the wall, etc., we can see that the tone colour of an instrument is determined by a very complicated mechanism.

The problem for my theory of sound is that it leads to a revisionary account of what we hear when we listen to an oboe performance. The sound radiated at the bell is a mixture of the wave arriving there the first time and those which have reflected multiple times in the bore. In other words, the total fluctuations outside the instrument, including those which cause our auditory experiences, contain multiple constituent fluctuations caused by different temporal parts of the event source at the reed. This means we hear different temporal parts of the event source simultaneously. Alternatively, it also means that the same temporal part of the event source is heard several times. However, our auditory experiences do not represent them as such. Our auditory system fails to separate the detected total fluctuations into multiple constituent fluctuations, and hence represent them as a single individual. Nonetheless, since it is the same event source producing all the constituent fluctuations we hear, it is still true that we hear one compression wave.

Moreover, the interface at the bell serves as a filter which shapes the spectrum of the reflected wave and the transmitted wave. Therefore, the total fluctuations both inside and outside the instrument should have a distorted spectrum compared to the spectrum of the event source. This implies that our perception of the event source is also distorted. Since the event source is partly determined by the reflected wave, this means the event source cannot be what it is without our perception of it being distorted. Our auditory experiences thus necessarily misrepresent the event source in this case.

It can now be seen that my view of sound leads to some surprising and perhaps undesirable implications regarding the case of the oboe (and other woodwinds). The complication of the account might also be a reason for objection. Understandably, a simpler account would be preferred. We would not have expected any need to differentiate the compression wave inside and outside the instrument into parts produced by different temporal parts of the vibration of the reed. However, such a differentiation follows directly from the identity condition of constituent fluctuation proposed in §6.8.3, which says that a constituent fluctuation cannot be caused by different temporal parts of its event source simultaneously. Giving up this condition would cost me the ability to explain the case of echo, especially those echoes which are distinguishable from but partly overlap with the corresponding primary sounds. It is because the overlapping parts are also caused by different temporal parts of the same event source.

We also would not have considered our perception of the vibration of the reed to be necessarily distorted. However, I suspect that the very same account might face less resistance if we instead describe the compression waves as "beautified" by the resonance. So, this might not be a serious problem after all.

My response to these problems is to insist on my metaphysical analysis and accept its implications on what we hear when we listen to an oboe performance. I wonder if there is any plausible reason to expect that we should be able to experience each of the overlapping parts of the compression wave separately. Experiencing the parts separately would be subjectively like hearing multiple echoes, which would be confusing and hinder the perception of the event source. Therefore, in some cases, the clarity of the perception of compression waves and that of the perception of event sources may be in conflict.

Besides, the distortion introduced by resonance needs not be bad. To see this, we should move over from the case of musical instruments to the case of our voice. The main difference between vowels is the frequencies of the formants, which are the harmonics strengthened by the resonances in the vocal tract. Resonance is, therefore, one of the many factors which make speech possible.[40] The case

[40] A kind of singing technique called overtone singing is based on the same principle in selectively amplifying one of the overtones. As a result, a singer can sing two notes at the same time: one at the

of resonance in human voice shows that although the main function of hearing is to acquire information about sound sources, how our perception of event sources is distorted can also be a source of useful information. It would, therefore, be less beneficial if our auditory system could distinguish all the overlapping constituent fluctuations, as it would vastly increase the number of individuals represented in our auditory experiences and make it very difficult to attend to the important ones at the same time. Not separating the constituent fluctuations may in some cases be the more efficient way of perception. Such a beneficial limitation of our hearing ability is what our ancestors unknowingly make use of in creating musical instruments. The art of musical instrument making is, therefore, the art of making tools for the creation of desirable distortion to our perception of sound sources.

6.9.4 Sound Reproduction

There seems to be (at least) two ways to understand “sound reproduction”. In one way, it means to produce a copy of a recorded sound. In another way, it means to bring the recorded sound back into existence. As a non-native speaker of English, the first one is my default understanding, and I was surprised to realise that there is the second understanding when I came across it in Martin (2012, p. 345). Anyway, in this subsection, I give an account of sound reproduction by first discussing Martin’s view. My view is that sounds as compression waves are concrete particulars which cannot be brought back into existence. What happens when we play an audio recording is that the playback event represents the auditory scenario captured in the recording. In other words, the case is parallel to the case of photography, where the captured scene is represented by an image. To this extent, the playback event is an auditory image of the original auditory scenario.

Sound reproduction can provide us with access to past events. How can this be explained? Martin (ibid.) claims that this can be explained only if the very sounds produced by the past events can be reproduced. This is because there cannot be non-

---

fundamental frequency and one at the amplified overtone. The more advanced polyphonic overtone singers can even sing two melodic lines simultaneously. According to my view, however, although we can hear two notes at once, there is only one sound, as there is only one event—the vibration of the vocal cords—which produces one compression wave the spectrum of which contains two peaks corresponding to the two notes we hear.

visual images (p. 344), such that the past events cannot be represented by any auditory image in a way parallel to how a photograph can be an image of the object captured. We shall then examine why there can only be visual images.

Martin holds that "a visual image presents the appearance of something which it does not exemplify" (p. 343). By itself, this does not explain why there cannot be auditory images. What he needs further is the sonicist idea that we hear non-sound entities by hearing sounds. According to him (p. 337), vision differs from the other senses in that we encounter both special objects of vision together with material objects as occupying the visual world. In contrast, sounds—the special objects of hearing—are merely *associated* with the material world. A mere *visibilia* like a hologram can present the appearance of a material object without exemplifying its material properties; but if a sound presents the appearance of having so-and-so auditory properties, it exemplifies those properties. As a result, no sound can be an auditory image.

Since we have accepted that sound sources can be directly heard, we can reject Martin's view by simply rejecting the sonicist assumption. Still, I need not then hold that we can hear mere *audibilia* alongside sound sources as occupying the auditory world. It is unclear whether there can be any mere *audibilia*—at least sounds as compression waves are not. Just like visual images need not be mere *visibilia*, there is no need for auditory images to be mere *audibilia*. All we need is a sound source which can appear to be something it is not.

The most common sound sources involved in sound reproduction nowadays are loudspeakers, and their job is to produce sounds similar to what you would hear in the presence of the recorded events. By playing back the recording, the vibrations of the speaker cones take on the auditory appearances of the recorded events without exemplifying it. We can hear the appearance of a glass shattering, but nothing is broken in the speaker. We can hear the appearance of a car going from left to right in the front, but the stereo system moved not an inch. In this sense, the vibratory events of the speaker cones are the auditory images which allow us to hear events in the past. This parallels the case of photography, where the pieces of paper on which the photos are printed serve as the visual images of the captured objects.

In Martin's view, holograms are the "limit case of nature images" (p. 342). When we look at a photograph, we are also aware of the piece of paper and some of its features, such as its flatness. In contrast, according to Martin, a perfect hologram can only be described in terms of the appearance of the represented objects (p. 340). However, no real hologram is perfect, so we can only imagine what such a thing would look like. For the present purpose, let us accept Martin's claim for the moment.

I would venture to say that loudspeakers are more similar to holograms than to photographs. It seems when we want to describe the vibration of speaker cones, we inevitably describe it in terms of the auditory appearances of the recorded events. Indeed, it would be a flaw of a loudspeaker if it adds features not belonging to the recorded events during playback. Of course, just like there is no perfect hologram, there is also no perfect loudspeaker. In reality, we normally have no difficulty in determining whether we are listening to an audio recording or not. Nonetheless, if we compare the case of audio playback with the case of photography, it seems loudspeakers are less prominent than the coloured pieces of paper in our experience, and therefore our attention is captured more fully by the recorded events in the auditory case than the captured objects in the visual case.

So far, I have been speaking generally about hearing as a form of event source perception. It might then be wondered what about the sounds produced? Can sounds be images as well? There seems to be two possibilities: a sound might be an image of another sound, or it might be an image of the event source of another sound. I concur with Martin in rejecting the first possibility. He is right that if a sound presents the appearance of another sound, it also exemplifies that appearance itself. Although in my view the two sounds have different properties such as different causal origins, such properties are not manifest in our experiences and hence are not parts of the appearances.

A sound might be a copy of another sound in a way similar to how wax figurines in Madame Tussauds are copies of celebrities as explained by Martin (p. 339), i.e. the originals and the copies share properties. In this sense, a sound can represent another sound not as an image but as a copy. However, this kind of representation is not a feature manifest in the appearance. It is rather something to be recognised based on the knowledge about the intended representational function for which the copy is produced (ibid.).

As for the second possibility, I leave it open. I have argued in §6.6.2 that treating sounds as representations of their event sources cannot explain in what sense sound sources can be objects of auditory experiences. However, this should allow that sounds might be images of their event sources and thereby represent them. If so, then a sound would present the appearance of its event source without exemplifying that appearance. If this event source is an image of another event source, the appearance presented by the sound would be exemplified not by its event source but by the recorded event, i.e. the event source of another sound.

Since there can be auditory images, the case of sound reproduction does not require us to treat sounds as entities which can be brought back into existence. We reproduce sounds by producing copies of them. The originals, however, are gone for good.

7 Conclusion

This dissertation has two aims.

The first aim is to defend wave theory. A part of this is done negatively in the criticism of other existing theories in Chapters 4 and 5, but mostly this is achieved in Chapter 6. The readers may be surprised by how few positive arguments I have put forward: only two, in §6.1 and §6.2 respectively. I also replied to two objections against wave theory in §6.3 and §6.4 and rejected an alternative theory in §6.7. In the rest of the chapter, I have done no more than fleshing out the details of my theory.

It is unclear how much more support to wave theory can be provided by philosophy. I believe the strongest support for wave theory comes from science. It is the theory's empirical success which leads to its wide acceptance in the scientific community and the general public. The evidence for it accrues as scientists painstakingly explain all sorts of sonic phenomena in terms of the behaviour of compression waves. If you want to know more about such evidence, my advice is to read an introduction to acoustics rather than a philosophical work like this one.

That said, philosophers do have their contribution. One example is the metaphysics of compression waves. This is philosophy proper. A view on this issue, such as mine which is presented in §6.5, can at best be science-informed. A complete scientific answer to such an ontological question is not forthcoming—partly because of the limit of the scientific method, and partly because scientists would probably be indifferent on this issue.

Another way philosophers can contribute to the debate is, perhaps unsurprisingly, to expose the problems of mistaken philosophical theories of sound. It might sound trivial to bring this up, but I believe the most fruitful lessons we can learn from the philosophy of sound are to be found in this negative enterprise. This leads to the second aim of this work: to show that, for wave theorists, the philosophy of sound is more about auditory perception than about sound. To my mind, this is the more important aim.

To this second aim, there is no direct discussion or argument up till now. However, this aim is what shapes the structure of the entire work. I first explained in Chapter 1 how the applicability of the veridicality requirement as a methodological assumption in the philosophy of sound depends on how we understood the phenomenology of auditory experience. More specifically, characterising an auditory experience as about a sound source instead of a sound would make it questionable what conclusion about sounds can be drawn from such an experience even if its veridicality is granted. Given the general practice of basing theories of sound on phenomenological claims about auditory experiences, the philosophy of sound in its current form may be problematic deep in its root.

This worry was further developed in Chapter 2 with a criticism of sonicism, by which I made a stronger case for the idea that what people take to be sounds in their auditory experiences may instead be the sound sources. Indeed, my conclusion was that it is more reasonable to take that which appears to be distally located to be the sound source instead. Chapters 1 and 2 together open up the possibility that the disagreement between theories of sound might steam from different theories of auditory perception in the background.

The claim that that which appears to be distally located in an auditory experience is the sound source itself has far-reaching consequences in the field of the philosophy of sound. I explored two main consequences in the following chapters.

First, Chapter 3 examined eight widely accepted features of sounds, four of which were shown to be questionable because they may instead be features of sound sources. Therefore, it is unjustified to argue from the appearance of something as processing these features to the conclusion that sounds should be identified with entities with these features.

Second, Chapters 4 and 5 continued with a survey of existing theories of sound. To the second aim of this work, the important lesson from this survey is not that these theories should be rejected—this is important for the first aim. Rather, it is that some of them should be rejected because of their reliance on the questionable characterisation of auditory experiences as being about sounds, when it is more reasonable to characterise them as being about sound sources instead.

For instance, all versions of distal theory could be considered as attempts to determine what exactly that which appears to be the distally located bearer of auditory properties is. By taking that distally located entities as sounds rather than sound sources, distal theorists turn the unproblematic claim that our auditory experiences do not represent compression waves as distally located into the denial that sounds are not compression waves.

Moreover, if O'Callaghan (2007b, p. 69) is representative here, distal theorists' denial that compression waves are heard can also be traced back to the same reason. As a sonicist, O'Callaghan holds that the direct auditory objects—that which appear as distally located—are sounds. Since compression waves are not heard in virtue of hearing those distal entities, compression waves are not heard. This denial, as I showed in §5.2.6, costs his theory the power to satisfactorily explain distinctive echo experiences. However, if he instead simply took the distal entities as the sound sources, his reason against hearing compression waves would no longer hold. He would then have the resources to better explain distinctive echo experiences.

Of course, distal theorists may suggest other reasons against hearing compression waves. If so, the tension between wave theory and distal theory cannot be completely resolved by simply reinterpreting our auditory experiences, as they still disagree on whether compression waves are represented. Nonetheless, the crucial observations of distal theory that auditory experiences represent distally located entities is perfectly consistent with wave theory. Indeed, to the extent that I accept that event sources are represented in auditory experience and identify event sources with disturbance events, there are probably more agreements than disagreements between my view and O'Callaghan's disturbance event theory.

If we put aside the identification statement of sound made by these theories and consider the phenomenological observations employed in support of them, we can easily see that philosophers of sound have taught us much about auditory perception indeed.

Here are some examples. The sound-less account of auditory perception by Young (2016) and the event source theory by Casati et al. (2013) forcefully challenge the conventional conception of our auditory access to the external world as mediated. The earlier view of Kulvicki (2008a, 2015) is a persuasive defence for the idea that we can hear properties of thing sources. His later idea of audible profiles in his (2017) not only highlights the perspectival aspect of auditory perception, but also expands the horizon of the auditorily accessible world to cover relational features of ordinary objects, events, and environs. Leddington (2014, 2019) compellingly shows that ordinary events are represented in auditory experiences as bearers of auditory properties. The distinction between disturbance events and transmission events drawn by O'Callaghan (2007b) constitutes a pressing request for a better account of the role of compression waves not only in auditory perception but more specifically in auditory experience. The active role of the auditory brain in shaping our auditory experiences as well as how this process is suited for the perception of sound sources are clearly explained by Nudds (2009, 2010b, 2015).

Such claims about auditory perception are all closely related to the theories of sound advanced by the respective authors. Although I rejected those theories, I think these claims are all invaluable contributions to our philosophical understanding of auditory perception. Indeed, all of them are parts of the background of my view presented in Chapter 6.

I believe the reason why claims about auditory perception are so crucial to the contemporary philosophy of sound is, as I mentioned already in §6.2, that the role as objects of hearing has been the most emphasised role of sounds in philosophical discussion. The inquiry into the nature of sounds thus becomes, or at least goes hand in hand with, the inquiry into what exactly are represented in auditory experiences. It is then unsurprising that so much has been learnt about auditory perception in the past twenty years.

Considered upon this background, the real challenge faced by wave theorists is not the one Sorensen (2009) tries to solve, namely to explain in what sense compression waves can be said to be located at their source. Rather, it is to provide a theory of auditory perception which (i) duly admits that distal entities—sound sources, reflective objects, etc.—are represented in auditory experiences simply as what they are rather than as some special auditory entities labelled as "sound" in the conventional conception; (ii) allows compression waves to be represented in auditory experiences; and (iii) explains in what sense compression waves are represented. If such a theory of auditory perception is available, the empirical success of wave theory will speak for itself. The aim of wave theorists as philosophers of sound is therefore not to improve our understanding of sounds as compression waves. Scientists are the more qualified people for this job. Rather, wave theorists should strive for a better understanding of auditory perception. To this extent, they are, or should more appropriately be considered as, philosophers of auditory perception instead.

References

Aristotle. (2001). *De anima* (J. A. Smith, Trans.). In R. McKeon (Ed.), *The basic works of Aristotle* (pp. 533-603). New York, NY: Modern Library.

Armstrong, D. M. (1961). *Perception and the physical world*. London, UK: Routledge & Kegan Paul.

Austin, J. L. (1962). *Sense and sensibilia*. Oxford, UK: Oxford University Press.

Baldwin, T. (1990). *G.E. Moore: The arguments of the philosophers*. London, UK: Routledge.

Berkeley, G. (1979). *Three dialogues between Hylas and Philonous* (R. M. Adams Ed.). Indianapolis, IN: Hackett Publishing Company. (Original work published 1713).

Bermúdez, J. L. (2000). Naturalized sense data. *Philosophy and Phenomenological Research, 61*(2), 353-374. doi:10.2307/2653655

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

Broad, C. D. (1952). Some elementary reflexions on sense-perception. *Philosophy: The Journal of the Royal Institute of Philosophy, 27*(100), 3-17.

Casati, R. (2018). Shadows, objects, and the lexicon: On some lexicalized and non-lexicalized concepts of shadow and light. In T. Crowther & C. Mac Cumhaill (Eds.), *Perceptual ephemera* (pp. 154-171). Oxford, UK and New York, NY: Oxford University Press.

Casati, R., Di Bona, E., & Dokic, J. (2013). The Ockhamization of the event sources of sound. *Analysis, 73*(3), 462-466. doi:10.1093/analys/ant035

Casati, R., & Dokic, J. (1994). *La philosophie du son*. Nîmes, France: Chambon.

Casati, R., & Dokic, J. (2009). Some varieties of spatial hearing. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 97-110). Oxford, UK: Oxford University Press.

Casati, R., & Dokic, J. (2010). Sounds. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2014 ed.). Retrieved from http://plato.stanford.edu/archives/fall2014/entries/sounds/.

Cohen, J. (2010). Sounds and temporality. In D. W. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. 5, pp. 303-320). Oxford, UK: Oxford University Press.

Crane, T. (2009). Intentionalism. In B. P. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford handbook of philosophy of mind* (pp. 474-493). Oxford, UK: Oxford University Press.

Fish, W. (2004). The direct/indirect distinction in contemporary philosophy of perception. *Essays in Philosophy: A Biannual Journal*, *5*(1), Article 5.

Foster, J. (2000). *The nature of perception*. Oxford, UK and New York, NY: Oxford University Press.

Gelfand, S. A. (2018). *Hearing: An introduction to psychological and physiological acoustics* (6th ed.). Boca Raton, FL: Taylor & Francis.

Heller, E. J. (2013). *Why you hear what you hear: An experiential approach to sound, music, and psychoacoustics*. Princeton, NJ: Princeton University Press.

Hilbert, D. R. (2005). Color constancy and the complexity of color. *Philosophical Topics, 33*(1), 141-158. doi:10.5840/philtopics20053315

Hopkins, R. (1998). *Picture, image and experience: A philosophical inquiry*. Cambridge, UK: Cambridge University Press.

Huemer, M. (2001). *Skepticism and the veil of perception*. Lanham, MD: Rowman & Littlefield.

Hume, D. (2007). *A treatise of human nature* (D. F. Norton & M. J. Norton Eds., Vol. 1). Oxford, UK and New York, NY: Oxford University Press. (Original work published 1739-1740).

Hyman, J. (2006). *The objective eye: Color, form, and reality in the theory of art*. Chicago, IL and London, UK: The University of Chicago Press.

Jackson, F. (1977). *Perception: A representative theory*. Cambridge, UK: Cambridge University Press.

Johansson, G. (1975). Visual motion perception. *Scientific American, 232*(6), 76-88. doi:10.1038/scientificamerican0675-76

Kalderon, M. E. (2018a). *The event of rarefaction: A defence and development of the wave theory of sound*. Manuscript in preparation.

Kalderon, M. E. (2018b). *Sympathy in perception*. Cambridge, UK: Cambridge University Press.

Kulvicki, J. (2008a). The nature of noise. *Philosophers' Imprint, 8*(11), 1-16.

Kulvicki, J. (2008b). [Review of the book *Sounds: A philosophical theory*, by C. O'Callaghan]. *Mind, 117*(468), 1112–1116. doi:10.1093/mind/fzn140

Kulvicki, J. (2015). Sound stimulants: Defending the stable disposition view. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 205-221). Oxford, UK: Oxford University Press.

Kulvicki, J. (2017). Auditory perspectives. In B. Nanay (Ed.), *Current controversies in philosophy of perception* (pp. 83-94). New York, NY: Routledge.

Leddington, J. (2014). What we hear. In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience* (pp. 321-334). Dordrecht, Netherlands: Springer.

Leddington, J. (2019). Sounds fully simplified. *Analysis, 79*(4), 621-629. doi:10.1093/analys/any075

Levin, J. (2000). Dispositional theories of color and the claims of common sense. *Philosophical Studies, 100*(2), 151–174. doi:10.1023/A:1018660204635

Lowe, E. J. (1981). Indirect perception and sense data. *The Philosophical Quarterly, 31*(125), 330-342. doi:10.2307/2219403

Maclachlan, D. C. L. (1989). *Philosophy of perception*. Englewood Cliffs, NJ: Prentice Hall.

Maclachlan, D. C. L. (2013). *The enigma of perception*. Montreal and Kingston, Canada: McGill-Queen's University Press.

Martin, M. G. F. (1992). Sight and touch. In T. Crane (Ed.), *The contents of experience: Essays on perception* (pp. 196-215). Cambridge, UK: Cambridge University Press.

Martin, M. G. F. (2007). Perception. In F. Jackson & M. Smith (Eds.), *The Oxford handbook of contemporary philosophy* (pp. 701-738). Oxford, UK: Oxford University Press.

Martin, M. G. F. (2012). Sounds and images. *British Journal of Aesthetics, 52*(4), 331-351. doi:10.1093/aesthj/ays036

Martin, M. G. F. (2017). Elusive objects. *Topoi, 36*(2), 247-271. doi:10.1007/s11245-016-9389-9

Matthen, M. (2010). On the diversity of auditory objects. *Review of Philosophy and Psychology, 1*(1), 63-89. doi:10.1007/s13164-009-0003-6

Matthiessen, H. O. (2010). Seeing and hearing directly. *Review of Philosophy and Psychology, 1*(1), 91-103. doi:10.1007/s13164-009-0005-4

Montague, M. (2016). *The given: Experience and its content*. Oxford, UK and New York, NY: Oxford University Press.

Moore, B. C. J. (2013). *An introduction to the psychology of hearing* (6th ed.). Leiden, Netherlands: Brill.

Mulligan, K., & Correia, F. (2017). Facts. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2017 ed.). Retrieved from https://plato.stanford.edu/archives/win2017/entries/facts/.

Nevid, J. S. (2016). *Essentials of psychology: Concepts and applications* (5th ed.). Boston, MA: Cengage Learning.

Nudds, M. (2001). Experiencing the production of sounds. *European Journal of Philosophy*, *9*(2), 210-229. doi:10.1111/1468-0378.00136

Nudds, M. (2009). Sounds and space. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 69-96). Oxford, UK: Oxford University Press.

Nudds, M. (2010a). What are auditory objects? *Review of Philosophy and Psychology*, *1*(1), 105-122. doi:10.1007/s13164-009-0003-6

Nudds, M. (2010b). What sounds are. In D. W. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. 5, pp. 279-302). Oxford, UK: Oxford University Press.

Nudds, M. (2014a). Auditory appearances. *Ratio: An International Journal of Analytic Philosophy, 27*(4), 462-482. doi:10.1111/rati.12083

Nudds, M. (2014b). Commentary on Leddington. In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience* (pp. 343-348). Dordrecht, Netherlands: Springer.

Nudds, M. (2015). Audition. In M. Matthen (Ed.), *The Oxford handbook of philosophy of perception* (pp. 274-293). Oxford, UK: Oxford University Press.

Nudds, M. (2017). The non-visual senses: The distinctive role of sounds and smell. In B. Nanay (Ed.), *Current controversies in philosophy of perception* (pp. 95-106). New York, NY: Routledge.

O'Callaghan, C. (2007a). Echoes. *The Monist, 90*(3), 403-414. doi:10.5840/monist200790328

O'Callaghan, C. (2007b). *Sounds: A philosophical theory*. Oxford, UK and New York, NY: Oxford University Press.

O'Callaghan, C. (2008). Seeing what you hear: Cross-modal illusions and perception. *Philosophical Issues, 18*(1), 316-338. doi:10.1111/j.1533-6077.2008.00150.x

O'Callaghan, C. (2009a). Audition. In J. Symons & P. Calvo (Eds.), *The Routledge companion to philosophy of psychology* (pp. 579-591). Abingdon, UK: Routledge.

O'Callaghan, C. (2009b). Sounds. In T. Bayne, A. Cleeremans, & P. Wilken (Eds.), *The Oxford companion to consciousness* (pp. 609-611). Oxford, UK: Oxford University Press.

O'Callaghan, C. (2009c). Sounds and events. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 26-49). Oxford, UK: Oxford University Press.

O'Callaghan, C. (2010a). Constructing a theory of sounds. In D. W. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. 5, pp. 247-270). Oxford, UK: Oxford University Press.

O'Callaghan, C. (2010b). Perceiving the locations of sounds. *Review of Philosophy and Psychology, 1*(1), 123-140. doi:10.1007/s13164-009-0001-8

O'Callaghan, C. (2011a). Hearing properties, effects or parts? *Proceedings of the Aristotelian Society, 111*(3), 375-405. doi:10.1111/j.1467-9264.2011.00315.x

O'Callaghan, C. (2011b). Lessons from beyond vision (sounds and audition). *Philosophical Studies, 153*(1), 143-160. doi:10.1007/s11098-010-9652-7

O'Callaghan, C., & Nudds, M. (2009). Introduction: The philosophy of sounds and auditory perception. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 1-25). Oxford, UK: Oxford University Press.

O'Shaughnessy, B. (2000). *Consciousness and the world*. Oxford, UK and New York, NY: Oxford University Press.

O'Shaughnessy, B. (2009). The location of a perceived sound. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 111-125). Oxford, UK: Oxford University Press.

Pasnau, R. (1999). What is sound? *The Philosophical Quarterly, 49*(196), 309-324. doi:10.1111/1467-9213.00144

Pasnau, R. (2009). The event of color. *Philosophical Studies, 142*(3), 353-369. doi:10.1007/s11098-007-9191-z

Perkins, M. (1983). *Sensing the world*. Indianapolis, IN: Hackett Publishing Company.

Roberts, P. (2017). Turning up the volume on the property view of sound. *Inquiry: An Interdisciplinary Journal of Philosophy, 60*(4), 337-357. doi:10.1080/0020174X.2016.1159979

Rossing, T. D. (2014). Introduction to acoustics. In T. D. Rossing (Ed.), *Springer handbook of acoustics* (2nd. ed., pp. 1-7). Heidelberg, Germany: Springer.

Schaffer, J. (2009). On what grounds what. In D. J. Chalmers, D. Manley, & R. Wasserman (Eds.), *Metametaphysics: New essays on the foundations of ontology* (pp. 347-383). Oxford, UK: Oxford University Press.

Schwitzgebel, E., & Gordon, M. S. (2011). Human echolocation. In E. Schwitzgebel (Ed.), *Perplexities of consciousness* (pp. 57-69). Cambridge, MA: The MIT Press.

Scruton, R. (1997). *The aesthetics of music*. Oxford, UK and New York, NY: Oxford University Press.

Scruton, R. (2009). Sounds as secondary objects and pure events. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 50-68). Oxford, UK: Oxford University Press.

Scruton, R. (2010). Hearing sounds. In D. W. Zimmerman (Ed.), *Oxford studies in metaphysics* (Vol. 5, pp. 271-278). Oxford, UK: Oxford University Press.

Sorensen, R. (2008). *Seeing dark thing: The philosophy of shadows*. New York, NY: Oxford University Press.

Sorensen, R. (2009). Hearing silence: The perception and introspection of absences. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays* (pp. 126-145). Oxford, UK: Oxford University Press.

Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. London, UK: Routledge.

Supa, M., Cotzin, M., & Dallenbach, K. M. (1944). "Facial vision": The perception of obstacles by the mind. *The American Journal of Psychology, 57*(2), 133-183. doi:10.2307/1416946

Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. Cambridge, MA and London, UK: The MIT Press.

Tye, M. (2009). *Consciousness revisited: Materialism without phenomenal concepts*. Cambridge, MA: The MIT Press.

von Wright, G. H. (2000). Some observations on sound and hearing. *SATS – Nordic Journal of Philosophy, 1*(1), 7-11. doi:10.1515/SATS.2000.7

Walton, K. L. (1984). Transparent pictures: On the nature of photographic realism. *Critical Inquiry, 11*(2), 246-277. doi:10.1086/448287

Young, N. (2016). *Not hearing sounds* (Doctoral dissertation). University of Antwerp, Antwerp, Belgium.

Young, N. (2017). Hearing spaces. *Australasian Journal of Philosophy, 95*(2), 242-255. doi:10.1080/00048402.2016.1164202

Young, N. (2018). Hearing objects and events. *Philosophical Studies, 175*(11), 2931-2950. doi:10.1007/s11098-017-0988-0