

Are there auditory objects in the auditory domain, like visual objects in the visual domain? If so, what is an auditory object? Discuss in the light of evidence from philosophical theories of auditory/object perception, empirical work in neuroscience, or both.

Sam Wilkinson¹
University of Edinburgh

One can understand the word “object” as a concrete physical object or as that which is on the receiving end of a subject-object relation, namely, that entity which a particular cognitive state or process is “of.” These latter objects are determined by the way our sensory systems exploit the ways elements of the world impinge upon our bodily surfaces. Our visual system exploits light reflected off the surfaces of objects; therefore, the objects of our visual experiences can be physical objects just sitting still. Our auditory system, on the other hand, exploits sound waves caused by impacts of one sort or another, and these are events. You don’t, strictly speaking, hear a billiard ball or a car: You hear a billiard ball hit another one, or a car drive past.

There are two reasons why one might be tempted to deny that there are auditory objects in anything like the way that there are visual objects. One is the naïve confusion between these two meanings of the word object: The objects of audition aren’t objects qua lumps of matter. The other, which is to be taken more seriously, is that a sound itself shouldn’t classify as an object of audition any more than visible light should classify an object of vision. It is, rather, thanks to sounds that we know about goings on in the world. And yet, when we are informed about the outside world through audition, this is a far more fallible and inferential mechanism than when we are so informed through vision. I think that this way of thinking should be heavily moderated. We can be, and often are, non-inferentially and reliably informed about the world through audition. The trouble is that the complexity and reliability of the mechanisms at play in audition are less obviously apparent to us than they are in the case of vision.

The fact that vision and audition exploit different environmental regularities means that, first, the type of information in terms of raw input is different and, second, that it gets processed differently. In the philosophical mainstream, this distinction between the raw input and the way in which it is processed is understood in terms of conceptualization. Instead of discussing this tricky notion, I will cash things out in terms of transparency. There are two different notions of transparency: One is a claim about whether, when we undergo a qualitative experience, we can be said to be aware of things in the world (and not just qualia or sense-data). The second notion, and the one that interests us here, is about the relationship between distal and proximal stimuli. A useful way of thinking about this kind of transparency is to consider sensory substitution (Bachy-Rita, 1967). Sightless people can “see” thanks to information coming in through a camera on their head, which gets transferred into tactile stimulation on their back (or tongue). They begin by attending to the proximal stimulus (what they are feeling on their back or tongue) and form explicit inferences about the distal stimuli that may have caused it. After habituation, however,

¹ s.l.j.wilkinson@sms.ed.ac.uk

something fascinating happens: They start to attend to the distal stimulus directly, and with this comes a radical phenomenological shift. The subjects now feel as though this is something close to seeing. Furthermore, judgments about the world are no longer based on explicit rational inference.

What it is for something to be the object of a sensory experience, I think, has a lot to do with this notion of attention and attentional shift. It is a precondition of our attention being on the distal stimulus that our experience commits us to the existence of that distal stimulus, where commitment is understood as the forming of judgments (with accuracy conditions) about the presence and properties of things in the world. Just as our senses generally commit us to the existence of things in the world, so it is that our auditory experiences in particular commit us to the existence of things other than sounds: namely, the events (and objects implicated in such events) that produce sounds. This can be understood in terms of proper function: It is the function of the senses to inform us about the world, and audition is no different. One consequence of this view is that our auditory system may perform its function properly at the lower level of veridically representing the sound, but malfunction at the higher level of wrongly representing the source of that sound.

Although we have sensory systems, which exploit environmental regularities detecting and processing them in such a way as to inform us, qua organisms we benefit from this totally independent of our ability to appreciate the underlying mechanisms. If these mechanisms are, as a matter of fact, intricate and reliable in informing us, however, then it clearly indicates that we, as organisms, are justified in relying on audition to tell us about things in the world. We, as philosophical theoreticians, are also justified in saying that there are auditory objects and that these are not merely sounds but worldly events. In other words, if audition of a healthy everyday sort (not based on explicit reference) is transparent in such a way as to commit us to the existence of worldly entities (including events), then it is quite simply these entities which are the proper objects of audition.

Humans are very good at recognizing the sources of sounds: When you wake up in the night to the sound of footsteps, it is hardly something that is the product of a deliberate inference: You hear it straight off as footsteps. We are also very good at perceiving the properties of these sources. It has been established, for example, that we are good at hearing the trajectory of an approaching sound source (Neuhoff, 2004), at judging the time at which we will collide with it (Schiff and Oldak, 1990), and at judging whether a sound source is within reach. These sorts of egocentric judgments are clearly highly adaptive in obstacle and projectile avoidance. It is, therefore, more surprising that more allocentric judgments can also be formed on the basis of audition: For example, when flat steel plates of different shapes (circular, square, and triangular) are struck, listeners are able to classify the shapes at a level well above chance (Kunkler-Peck and Turvey, 2000). How is this done, and why do our own abilities surprise us?

The auditory system rules certain things out statistically: It is unlikely that a sound with a given frequency, timbre, and a certain progression of decay could come from anything other than an event of a certain kind. As Nudds (forthcoming) puts it, “the particular pattern of frequency components produced by a material object when it vibrates is determined in a law-like way by both the physical nature of the object and the nature of the event that caused it to vibrate.” Note

how I say that it is “the auditory system” that must rule certain things out. This is because these judgments (if they can be called that) are not at the personal level; they are not the products of rational deliberation, and relatedly, their justification is harder to express. Indeed, with the task involving the steel plates, the resulting judgments aren’t even accompanied by a sense of justification (Kunkler-Peck and Turvey, 2000).²

One reason why we are more aware of our visual capacities than our auditory ones is that vision furnishes us with more input for thought and reasoning. Visual percepts are more frequently and variably conceptualized, by which I intend to commit myself to saying no more than that they serve as input for systematic inferential (and self-evaluative/metacognitive) processes at a personal level. (Perhaps this is due to the intrinsically spatial and geometric nature of visual representation.³) That is why we are more capable of justifying judgments made on the basis of what we see than on the basis of what we hear. But this is not to say that audition is not typically transparent and reliable.

We speak of hearing not just the things which produce sounds, but sounds themselves. For example, when we take a “musical stance,” we attend to the sounds and don’t form (practical) judgments about the sources of the sounds. And we don’t need to: The bassoonist, one hopes, is not going to throw his instrument at you.⁴ But music is an exception among the class of auditory experiences. Like other senses, audition informs us about the world. The discriminatory mechanisms that have enabled this, which we use in the “practical stance,” have been subsequently used, and possibly further developed, in the culturally-driven appreciation of music (and relatedly, vocal interaction), but this is not the proper function of audition; namely, it is not the evolutionary reason for why we hear at all. Similarly, it is not sounds but worldly events that are the proper objects of audition.

References

- Bach-y-Rita, P. (1967). Sensory Plasticity. *Acta Neurologica Scandinavica*, 43, 417-26.
- Kunkler-Peck, A., & M.T. Turvey. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 279–294.
- Neuhoff, J. (2004). Auditory motion and localization. In J. Neuhoff (Ed.), *Ecological Acoustics*. London: Academic.
- Nudds, M. (in press) What are auditory objects? *Review of Philosophy and Psychology*.
- Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 303–316.

² Note how this feeling of guess-work is reminiscent of blindsight patients (Weiskrantz 1986).

³ Also, vision is obviously informationally more complex: the retina is more discriminating than the ear drum, in part since the photon is smaller than the air molecule.

⁴That is in part why certain forms of electronic music, with purely synthesized and unrecognizable sound sources work fine as music: People generally aren’t frustrated that they can’t recognize the sound source since they don’t need to.