Still My Guitar Gently Weeps.
Questions for an Ockhamized Metaphysics of the Event Sources of Sound

LUCA GASPARRI

Abstract
Casati, Di Bona and Dokic have recently argued that sounds are identical to their event sources. In this paper, I review the arguments they have offered in support of this view, show that their claims fail to defend it in a completely persuasive and conclusive fashion, and present some new questions for their thesis.

1 What Are Sounds?
Ordinary parlance is used to distinguishing sounds from the objects and the events that produce them. If I strike the fifth string of a guitar with my fingernail and play an A, I can separate three different things: the vibration of the string, the series of mechanical waves propagating through air as a result of the vibration of the string, and the subjective experience or sensation of the A. Which of these three elements is the sound? Answering this question amounts to favoring a particular competitor in the dispute among proximal, medial and distal accounts of the relation between sounds and event sources (cf. section 2). According to proximal theories, the sound is the subjective experience caused by the vibration of the string through the series of mechanical waves it generates in the environment; according to medial theories, the sound is the series of mechanical waves generated in the environment by the vibration of the string; according to distal theories, the sound is the vibration of the string. Casey O’Callaghan (cf. [20]) has defended a version of the distal thesis enriched by a distinction between distal sounds and audible events. In his view, also called the mereological or parthood account, the collision between the fingernail and the string gives rise to two events: the vibration the string (the event source) and the sound of the string.
The two are both located at the distal level, but are distinct and connected by a mereological relation such that the sound is always a proper part of its event source.

2 Ockhamizing the Parthood Account

In a recent, very nicely argued and thought-provoking paper, Roberto Casati, Elvira Di Bona and Jérôme Dokic have suggested that the parthood account should be improved by means of an ockhamization of its central mereological claim. According to their thesis, called the identity view, “the sound is not a proper part of a distinct event that is its source; it is identical with […] the event source” ([2, p.464]). In terms of our example, the vibration of the fifth string does not entertain a whole-part relation with the sound it produces: rather, it is that sound. Since there is no difference between hearing sounds and hearing event sources, there is no difference between hearing the vibration of the string and hearing the sound it allegedly generates, and talk of ‘sounds’ is entirely redundant on talk of ‘event sources’: we “hear sounds” because sounds are distal vibratory events that we can access in auditory modality.

In this note, I shall not enter into the intricate landscape of the competition between proximal, medial and distal accounts of the relation between sounds and event sources, nor will I be directly concerned with the general issue of determining whether espousing a distal framework is the best way to go if one wants to set up a solid metaphysics of sound (cf. [18], [23]). Rather, I will restrict myself to pointing out the reasons why I think that the ockhamization advocated by Casati et al. does not return a completely attractive development of the parthood account. In doing this, I shall not assume or maintain that the parthood account is the best theory of sound on the market. I will simply argue that the position endorsed by O’Callaghan, once ockhamized, does not yield a fully convincing thesis about the nature of sounds, while remaining neutral on whether the original, non-ockhamized version of the parthood account is overall a robust and appealing theory.

Casati et al. argue that the identity view is just an innocent incremental restriction of the mereological account, since identity is the limit case of parthood. In addition, they observe that although the identity view requires its endorser to give up ordinary language statements distinguishing sounds and event sources, such as “I heard the collision of the fingernail with the string and the attack of the A”, this cost is readily counterbalanced by the “natural language propension to classify the
objects of auditory perception in terms of events, with no reference to a supplementary category of separate sounds” ([2, p.464]). As a matter of fact, when presented with the recordings of auditory events, listeners tend to identify sounds in terms of the happenings which caused them, focusing on their proximal attributes or on their object source only when they fail to recognize what distal event is responsible for their occurrence (e.g., cf. [7], [17], [28]). As James J. Gibson noted, “physical acoustics tells the man in the street that sensations of loudness, pitch, and pitch mixture are in his head, and only arise because they correspond to the variables of sound waves in the air. […] Nevertheless he goes on hearing natural events like rubbing, scraping, rolling, and brushing, or vocal events like growling, barking, singing, and croaking” ([8, p.170]). So we do have a somewhat natural inclination to depict the phenomenology of hearing as if we could have direct auditory access to distal events and “hear vibrations” (cf. [10], [21]). But is this really that relevant when it comes to corroborating the strong metaphysical claim the identity view wants us to accept?

3 Some Problems

Let me start with confessing that I fail to see the point of appealing to the way ordinary talk frames auditory perception as a means to assess the attractiveness of the identity view. First of all, language use licenses conflicting modes of reference to sounds, and the “natural language propension” recalled by Casati et al. is far from systematic or free of major exceptions. When a listener reports that she has “heard a guitar” or “heard a loud buzz”, we do not feel irresistibly pushed to stigmatize her linguistic behavior as deviant or uncooperative. In addition, while our ordinary talk of sounds is often in tune with the distal position, our ordinary thinking of sounds seems to be implicitly medial. For example, when presented with the sound of a collision, we talk of “hearing the collision”, but we think of “becoming aware of the collision thanks to our auditory access to the sound waves it has introduced in the environment” (cf. [23]). Finally (and most importantly), the identity view is a case against folk acoustics. One of its distinctive assumptions is that the received habit of locating sounds in the hearer (or in the medium between the resonating object and the hearer) is biased, referentially misplaced, haunted by misleading metaphors, and that it should play no positive role in a mature metaphysics of sound. These being the premises, either one rejects the identity view and accepts that folk acoustics can be used
as a platform where to test the strength of a given metaphysics of sound, or one accepts that the identification of sounds with their event sources is viable despite its inconsistency with the layman’s view of sound. In the latter case, pre-theoretical propensities need to be banned from the argumentation arena.

Second, it seems that some of the properties that ordinary parlance ascribes to sounds cannot be predicated of event sources without generating unwelcome consequences. Casati et al. maintain that the problem is easy to deal with, since the only properties of event and object sources that resist being attributed to sounds are the non-auditory ones (e.g., the release of energy triggered by the vibration of a string, as contrasted to pitch, loudness, or timbre). Which means that the identification of sounds with event sources can be secured on grounds that the set comprising the auditory property-types that can be predicated of sounds but not of event sources is prima facie empty. Yet, it is hard not to find this claim problematic. For the sake of our argument, let us allow that sounds and sequences of sounds are things that can be heard, that can propagate through space, and that can be the object of a structural or qualitative evaluation (as it happens, e.g., when we identify a mistake in a chord progression or characterize a musical piece as ‘relaxing’, ‘solemn’, or ‘melancholic’). Idealizing a lot: for every $x$, if $x$ is a sound, $x$ can be heard, $x$ can propagate through space, and sequences of $x$-type items can be the object of a structural or qualitative evaluation. Now assign the variable $x$ the value event sources. Complaining about the plausibility of “vibratory event sources can be heard” would be question begging, since the metaphysical tenability of this postulate is precisely what the identity view aims to defend. But what about the other two cases?

4 Fuzzy Boundaries

In the former case, it seems that in order to reconcile the identity view with the intuition that sounds propagate through space, we must ascribe spatial diffusion to entities that are inherently local. If sounds “fill the space” and “come to our ears”, and sounds are distal vibratory events, then distal vibratory events “fill the space” and “come to our ears”? Of course not. Space is filled by the oscillations of pressure introduced into the atmosphere by vibrating objects (e.g., resonating forks), changes in the surface of a liquid body (e.g., water splashes), and local aerodynamic events (e.g., exploding balloons), not by vibrating objects, liquid bodies,
and local aerodynamic events themselves (cf. [6]). So all the worse for the intuition that sounds can propagate through space? Proponents of the picture defended by Casati et al. are likely to insist that this is exactly the point the identity view is trying to make, since one of the purposes of their ockhamization is to deny that allegedly meaningful idioms such as “sounds fill rooms” describe substantial facts about sounds. In Casati and Dokic’s words, “one does not create sounds by surrounding vibrating objects with a medium – one simply reveals them” ([3]). But some complications arise.

Suppose that my guitar is resting on my left leg. After the collision with my fingernail, its fifth string starts vibrating. An infinitesimal time fraction later, so will do the headstock, the neck, the heel, the soundboard, the fretboard, and the other five strings. Harmonic resonance and the mechanical response of the body of the guitar will affect the timbre, the volume and the sustain of the A. In turn, the vibration of the entire guitar will be backward influenced by the way it is absorbed and retransmitted to the instrument by my jeans, by the organic material of my leg, by the surface of my chair, and so forth (cf. [13]). The example poses two problems. First, if the vibratory event generating the A is spatially spread in this macrosystem, then we are still dealing with a theory allowing that sounds, qua vibratory events, can spread through space. Second, how are we supposed to localize the vibratory event ‘A’ in this scenario and, most importantly, how can we determine its upper boundaries? Is there any substantial, not just allusive reason why we should rule out an unbuttoned spatial incrementation of the event source centered on the vibration of the string and deem impossible to maintain that the portion of space occupied by the occurrence of the event ‘A’ comprises the whole environment where the A gives rise to a vibration, up to and including the mechanical response of the organs responsible for neural transduction in the inner ear of the listener (e.g., the vibration of the basilar membrane in the cochlea) (cf. [16], [24])? Which means: how exactly do we preserve the distality of the theory after its ockhamization?

Building on an analogy with the visual domain, the advocate of the identity view might want to respond that the event source ‘A’ can be confined at the distal level by observing that the properties of the vehicle thanks to which the listener is able to hear the sound of the guitar (in this case, the density, the temperature and the viscosity of air) influence the perception of the A just as the intensity and the type of light in a room shape our perception of colored surfaces. The reply could thus run
as follows: since elastic intermediaries affect the mode of presentation of distal audible events without altering their constitutive properties, the events occurring between the vibration of the string and the listener can be safely characterized as medial and kept out of the spatial boundaries of the event source because they constrain the way the listener auditorily accesses the A, but not the A itself. Yet, it is hard to tell how this proposition could remain attractive in addressing cases where the relevant event sources are not centered on the vibration of a solid object, but on the oscillation of air itself, as it happens with thunders or musical instruments like flutes (cf. [3]). Furthermore, suppose that I am playing my guitar with my chin leaning against its side. I pluck the fifth string. Suddenly, I feel my skull vibrating, and I notice that the A “sounds” sharply different from the way it “sounded” to me while I was playing the instrument in the normal position. Now, assuming that the A I am hearing is nothing but a mechanical vibration, how are we supposed to pin down an inventory of the objects whose vibration contributes to the constitution of the event ‘A’ in this scenario? Am I hearing my skull? Am I hearing the string and the body of the guitar through my skull? Am I hearing the string through my skull and the body of the guitar? Am I hearing the A through two separate vibratory channels, one flowing from the string to my ears through air, one flowing from the string to my ears through the body of the guitar and my skull? It seems there are too many open options here to maintain that the identity view simplifies our treatment of auditory experience. It is important to note that these difficulties are generated by the ockhamization proposed by Casati et al., because the disentanglement of event sources and sounds originally recommended by the parthood account (cf. [19], [20]) allowed to cut through the problem. In fact, if sound tokens are distinct from their vibratory source, it is not necessary to solve general metaphysical problems related to the fuzziness of vibratory events in order to guarantee sounds a clear-cut localization in the environment, provided that sounds must be spatially localized in some way (which is denied by some: e.g., cf. [5], [22], [25], [26], [27]).

5 Musical Grammars and Acoustic Qualia

Let me now turn to the structural and qualitative features of sequences of sounds. Assume the perspective of a framework like the generative theory of tonal music (GTTM). As the informed reader will recall, the rationale of GTTM is provided by the acknowledgment that experienced
listeners with little formal training in the tonal idiom are effortlessly able to identify a previously unknown piece as an example of the idiom, to recognize elements of a piece as typical or anomalous given the regularities of the idiom, to identify a performer’s error as possibly producing an ill-formed structure, and so forth (cf. [12], [14], [15]). Since this spectrum of capacities arises from listeners’ spontaneous analysis of tonal inputs into structured mental representations, it is possible to characterize the ability to pair strings of surface audible configurations with analytical representations in terms of the knowledge of a set of principles defining a “grammar”, and qualify intuitions of tonal well-formedness as the correlates of an implicit knowledge of the “grammatical” structures that are possible or admissible in the tonal idiom (cf. [4]). Now, suppose that a given listener $L$ is presented with a tonal sequence $K$ and is asked to express her opinion regarding the structural well-formedness of $K$. Suppose also that, in reply, $L$ states “$K$ is a well-formed musical configuration”. Now apply the identity view and replace $K$ with “the sequence of distal vibrations underlying $K$” in the response given by $L$. Two obvious problems arise. First, there is more than one reason to expect that the sentence “the sequence of distal vibrations underlying $K$ is a well-formed musical configuration” will not be judged by $L$ as a consistent means to express her belief about the well-formedness of the sequence of sounds she has just heard. Second, there is room to doubt that “the sequence of distal vibrations underlying $K$ is a well-formed musical configuration” would be a sentence making any sense at all. One may be tempted to accommodate the objection by devising a procedure to systematically correlate the predicates of $K$ computed by $L$ in evaluating whether $K$ is an example of the tonal idiom to predicates in sounding objects, but the situation would not change much: it still seems counterintuitive and unnecessarily reductive to say that what is analyzed by $L$ in the evaluation of $K$ is a sequence of distal vibratory events rather than a (possibly distal) acoustic surface distinct from its vibratory source.

In a similar vein, one might object that if sounds were vibratory event sources, we should not feel excessively uncomfortable in equating the object of our judgments about the qualitative aspects of musical sequences with their distal source. Again, this is debatable. Right now I am sitting in front of my computer, with headphones on my ears, and I am listening to Julian Bream’s interpretation of *La Maja de Goya* by Enrique Granados, for solo classical guitar. If I say that the *tonadilla* I am hearing is soft, amusing, and sensual, what I point out is not that
the sequence of vibrations enacted by the membranes of my headphones is soft, amusing, and sensual, but that the sound they produce exemplifies those qualities in my personal experience of the piece. It is true that the identity view is not bound to deny that listeners can identify aesthetic or emotional qualia in their subjective experience of music and describe them via hedonic terms (cf. [11]). Yet the problem seems to remain. Consider a listener \( L \) expressing a judgment about the aesthetic attributes of a sequence of sounds \( K \): will she accept that her assertion can be charitably interpreted as attaching qualitative predicates to a series of mechanical vibrations? Hardly so. Plus, \( L \) will probably protest that her characterization of \( K \) as ‘soft’ or ‘sensual’ is not intended to maintain that her experience of \( K \) is soft or sensual, but that \( K \) itself has those properties, and that her ability to experience a sense of softness or sensuality as a result of her exposure to \( K \) is a function of \( K \)’s “having softness and sensuality”. In this respect, the identity view seems to collide with the essentially acousmatic character of phenomenology of music: we tend to perceive music as detached from the physical circumstances of its production (cf. [9], [26]).

Let me conclude with one last observation. If I record the sound made by the fifth string of a guitar with a lossy compression algorithm, I transfer it on my computer and I listen to it again on my headphones, it seems quite improper to say that what I am replicating is a distal vibratory event. Proponents of the identity view may wish to reply that this is not a conclusive objection, since after all what I am replicating is just an acoustic surface and the feasibility of this procedure is consistent with a plurality of accounts of the constitution of sound tokens. Generally speaking, one would think that if some property \( P \) is constitutive of an entity \( x \), then \( P \) must explain why \( x \) looks the way it does, and I have no intelligent complaint against the assumption that the properties of distal vibrations constitute sound tokens in this sense. The problem is that unless we agree on giving up the intuition that sound tokens can be recorded (or heard), the attributes of distal vibrations cannot be the only constitutive properties of sound tokens, because if the attributes of distal vibrations were the only constitutive properties of sound tokens, then the only things that could be directly recorded (or heard) would be medially-shaped modes of presentations of source vibrations, rather than ‘sounds’ in the sense the expression assumes within the identity view.
6 Conclusion

The identity view offers a very straightforward way to account for the sense of unity we experience when we perceive sounds and event sources, it furnishes immediate metaphysical grounds to our ability to perceive sounds as spatially located, and it remains elegantly neutral on the relation between object sources and event sources. Plus, none of the arguments supplied above provides a completely knock-out case against the identification of sounds with event sources: in all honesty, I fail to see a principled reason why the identity view should not be able to accommodate them in the long run. But in order to make the identification of sounds with distal vibratory events definitely worth the deal, we need to have a more distinct sense of its advantages, rather than of its mere viability. We now know that if we license the attribution of audible properties to event sources, we must do without the assumption that sounds and distal vibrations are very different sorts of things. The price to pay in terms of intuitions is clear. But what do we get in return? Do we get a well-behaved framework whose simplicity, ontological parsimony and explanatory power are superior enough to those of the non-ockhamized view not to make us regret giving up our old-fashioned distinction between sounds and events sources? For the moment, this is far less clear. I believe that proponents of the identity view could benefit from a deeper inquiry into the problems presented in this paper, and I hope to have provided them with some useful questions by whose investigation they might further develop their already stimulating thesis.¹
Notes

1 I am grateful to Elvira Di Bona, Guido Andreolli and Giovanni Alli for cheerful discussion and valuable input on the first draft of this paper. I also thank two anonymous referees of this journal for their helpful comments on the submitted manuscript.

Luca Gasparri
Center for Neurocognition and Theoretical Syntax (NeTS)
Institute for Advanced Study (IUSS)
Palazzo del Broletto, Piazza della Vittoria 15
27100 Pavia, Italy

<lcgasparri@gmail.com>
<http://lucagasparri.com/>
References


