Final version forthcoming in *Philosophia*

Recorded Sounds and Auditory Media

Vivian Mizrahi

> "Before the invention of sound-reproduction technologies, we are told, sound withered away. It existed only as it went out of existence. Once telephones, phonographs, and radios populated our world, sound had lost a little of its ephemeral character. The voice became a little more unmoored from the body, and people's ears could take them into the past or across vast distances." (Jonathan Sterne, The Audible Past: Cultural Origins of Sound Reproduction, p. 1)

Abstract

A widespread view among philosophers and scientists is that recorded sounds and assisted hearing differ fundamentally from natural sounds and direct hearing. It is commonly claimed, for example, that the sounds we hear over the phone are not sounds emitted by the voice of our interlocutor, but the sounds reproduced by the phone's loudspeaker. According to this view, hearing distant sounds through communication and audio equipment is at best indirect and at worst illusory. In what follows, I shall reject these claims and argue in favor of a transparent view of auditory media, including radio, telephone, phonograph, etc. According to this approach, the great gift of Scott de Martinville and Edison is not to have invented devices able to reproduce vanished sounds but rather to have created technological instruments literally able to store and transmit them to future and distant listeners.

1. Introduction

Sterne's elegant introductory remark to his historical monograph makes us realize that in our world, where phones are ubiquitous and where a recording of almost any musical work can be accessed in a matter of seconds, it is difficult to imagine what the audible world was like before Bell's, Marconi's, and Edison's inventions. It was a world confined to the present and the local—a world where music was exclusively live and where discussions were face-to-face only. The auditory world we live in now is incontestably richer, because we can not only listen to falling rain or chat with our friends over dinner, but also listen to a Shostakovich concerto while commuting on the subway or close a deal over the phone with an entrepreneur in Tokyo while enjoying a direct view of the Swiss Alps.

Our audible world is richer, but is it real? Most contemporary philosophers refuse to put direct hearing and hearing through audio equipment, like the telephone, radio, or phonograph, on an equal footing, arguing that sounds delivered by those technological means are somewhat illusory. According to Kulvicki, for example, hearing music through loudspeakers is illusory, because speakers are designed to reproduce sounds they don't have. He explains:

> A speaker can vibrate as a guitar would vibrate even though it doesn't have the sound of a guitar: an impressive piece of *trompe l'oreille*. One might object that we don't ordinarily take speakers to be generators of illusions, but practice suggests otherwise. As speakers multiply from stereophonic to "surround", we expect nothing less than the illusion not only that some object has a sound it does not really have but that we are hearing the sound of an object where there is in fact no object. Speakers sound like very little for the same reason that the screen in a movie theater is not painted bright red. Watching a film, we are as uninterested in the sounds of the speakers as we are in the silver of the screen. It's the impressions that matter. (Kulvicki, 2008: 6)

A similar view is defended in O'Callaghan (2007: 161):

> Hearing a recording furnishes a limited form of perceptual access to the performance itself mediated by your awareness of the current sound. In addition to being mediated and indirect, in virtue of the process of recording and reproducing the sound, your awareness of the original performance also is both impoverished and illusory along a number of dimensions.

There are obviously differences between hearing a live concert and hearing the same performance over the radio, but I don't believe there is anything illusory in hearing sounds through technological media like radio or audio recordings. In fact, I suggest that the tendency among philosophers to regard those perceptual experiences as illusory is based on a misconception regarding auditory media in particular and perceptual media in general. The goal of this paper is to correct this view by providing a detailed account of the nature of auditory media.

2. The nature of sounds

The world of sounds is diverse. There are splashes, rustles, rattles, twangs, clinks, exclamations, music, and so on. Those sounds inform us about objects, materials, events, and processes and often trigger intense emotional responses. Despite their constant presence in our lives, the topic of sounds has been neglected by philosophers for decades. And it is only recently that sounds and audition have emerged in the philosophical landscape and appeared in numerous and interesting debates. At the center of these discussions is the question of the nature of sounds. Interestingly enough, philosophers have provided very different theories of the nature of sounds and classify sounds in a variety of categories. Philosophers have argued that sounds are properties,[1] particulars,[2] and also abstract individuals.[3] Although the debate about the nature of sounds is rich

[1] Pasnau 1999, Kulvicki 2008, 2014, Leddington 2014, 2018
[2] Casati & Dokic 1994, 2010, O'Callaghan 2007
[3] Nudds 2010

and fascinating, this paper will leave aside questions concerning the ontological classification of sounds. I will instead follow Casati and Dokic's preference for a classification of theories of sounds based on the way those theories locate sounds. They explain:

> Indeed, the various philosophical pronouncements about the nature of sounds can be rather neatly classified according to the spatial status each of them assigns to sounds. Where are sounds? Are they anywhere? The main relevant families of answers include proximal, medial, distal, and aspatial theories. Proximal theories would claim that sounds are where the hearer is. Medial theories – exemplified by mainstream acoustics – locate sounds in the medium between the resonating object and the hearer. Distal theories consider sounds to be located at the resonating object. Finally, aspatial theories deny spatial relevance to sounds. There are significant variants of each of these. (Casati & Dokic 2010)

Although this paper will not presuppose much about the nature of sounds, I will favor a distal approach to sounds. As stressed by many philosophers, distal theories of sounds are grounded in the phenomenology of hearing. Except for some unusual sounds like tinnitus, which are experienced as located somewhere in the head of the hearer, most sounds are experienced as external and as located at a distance from the perceiver. If you happened to visit a Swiss village, you would soon be able to localize the village's church through the ringing of its bells. When hearing the church bells, villagers perceive not only the pitch of the bells, but also their location and direction. Unlike the nonspatial and proximal view of sounds, I will presuppose in this paper that sounds usually have stable and distal locations.[4] Yet, if sounds are located at a distance from the hearer, there must be some way for the auditory system to be causally related to the sounds. Unlike the tactile and the taste system, which appear to be in direct contact with their perceived objects, hearing requires a medium that acts as a causal intermediary between the sounds and the auditory system of the perceiver. To travel from their source to the ears of the listener, sounds need to travel

[4] Sounds can travel in space (i.e. ambulance) but contrarily to some medial theories they do not *always* travel.

through a material medium like air or water. The propagation of sounds is in effect a repetitive disturbance of a medium's particles. Once the first particle of the medium is set in motion by the disturbance of a vibrating object (the source of the sound), the sound wave is propagated through the air by means of a chain of particle-to-particle interactions. Although hearing relies on a medium to causally transmit sounds from its source to the ears of the listener, I will argue that air, like all perceptual media, is not audible. Unlike medial theories of sounds, which consider sounds to be features of the medium itself, I maintain that the auditory medium is transparent and imperceptible. As I will show, this point is crucial when considering the nature of recorded sounds, because listening to recorded sounds is not different from "direct" hearing, except for the medium through which sounds is carried to the listener.

3. The nature of auditory media

Although the notion of perceptual media emerged in Aristotle's philosophy as a central part of his attempt to explain perception, its resonance throughout the history of the philosophy of perception has surprisingly been quite muted. Aristotle recognized that the remoteness of perceptual objects from the perceiver requires the existence of a causal intermediary. He claimed in particular that the colors of distant objects could not be perceived if there was not a suitable medium capable of acting directly on the organ of sight. This idea is clearly expressed in *De Anima* (ii 7 418b13–22; Smith in Barnes 1984b, 33–34):

> Colour sets in movement what is transparent, e.g. the air, and that, extending continuously from the object of the organ, sets the latter in movement. Democritus misrepresents the facts when he expresses the opinion that if the interspace were empty one could distinctly see an ant on the vault of the sky; that is an impossibility. Seeing is due to an affection or change of what has the perceptive faculty, and it cannot be affected by the seen colour itself; it remains that it must be affected by what comes between. Hence it is indispensable that there be something in between—if there were nothing, so far from seeing with greater distinctness, we should see nothing at all.

Although the interspace between the perceived object and the observer may appear to be empty, Aristotle explains that perception at a distance involves the existence of a physical medium between the remote object being perceived and the observer. According to Aristotle, perception at a distance is possible insofar as there is a causal intermediary connecting the remote object to the sense organ. In this respect, Aristotle contrasts distal senses, like sight and audition, which rely on an external medium, to proximal senses, like taste and touch, which require bodily contact with the perceived object.

As finely observed by Kalderon (2015), Aristotle's view of perceptual media offers a resolution to Empedocles' puzzlement over the perception of remote objects. In short, Empedocles' principle states that to be perceptible is to be palpable to a sense organ. But this principle seems to conflict with the perception of remote objects by sight and audition, in which objects are too distant for the perceiver's sense organs to be in direct contact with them, as required by Empedocles' tactile view of the senses. As stressed by Kalderon, by rejecting the Empedoclean principle, Aristotle commits himself to offering an alternative view of perception capable of explaining how perceptual objects can be experienced while remaining at a distance from the sense organs. Aristotle's insight is to link the distal aspect of vision and audition to the existence of an external medium. Aristotle argues that unlike touch and taste, the perception of auditory and visual objects requires the intervention of an external medium.

Aristotle's theory of perception at a distance requires that the remote object be able to reach the sense organs by an intervening medium, but it also requires that the medium be transparent. Although a proper interpretation of Aristotle's view of transparency should be left to specialists (see Kalderon 2018), the role of transparency in his account of perceptual media seems quite direct. If the external medium causally mediates the relation between the perceived object and the sense organs, it cannot interfere and therefore cannot be perceived per se. For a substance or object to

work as a medium, it must be transparent in order to allow other objects to be perceived *through* it. In fact, as stressed by Aristotle, perceptual media cannot be directly perceived:

> Now there clearly is something which is transparent, and by "transparent" I mean, what is visible, and yet not visible in itself, but rather owing its visibility to the colour of something else; of this character are air, water, and many solid bodies.
> (Aristotle, *De Anima*, trans. J.A. Smith, II, 7)

This point is fundamental when considering the nature of sounds and the role of sound waves. Because sound waves transmitted by the medium are necessary for hearing, some philosophers identify sounds with the vibrations that travel through the air or any other acoustic medium. According to this approach, the sounds we hear must be distinguished from the *sources* of the sound, which correspond to the vibrating objects that cause the medium to vibrate. This *medial* approach to the nature of sounds therefore conflicts with the Aristotelian view of perceptual media, which considers the medium transparent and only indirectly perceptible. In fact, if sounds are sound waves, as claimed by the medial approach, there is no auditory medium in Aristotle's sense at all, because the mechanical vibrations occurring in the air or water are the proper object of auditory perception. Unlike Aristotle's theory, the wave theory of sounds does not consider audition distal and can therefore dispense with the existence of an external medium.[5]

The variety of acoustic media is incredibly rich, because almost any solid, liquid, or gas, and even plasma, can transmit sound waves. Although the stiffness and density of a material influence the speed at which sound waves travel, almost any material can transmit sound and therefore serve as an acoustic medium. The diversity of acoustic media explains why complete silence is extremely

[5] The goal of this paper is not to defend a distal theory of sounds against the medial approach. As asserted above, I will presuppose that the distal approach to sound is correct and explore its consequences for a theory of recorded sounds. It is interesting, however, to note that medial theories of sounds, by identifying sounds with sound waves, cannot theoretically distinguish between hearing a sound directly and hearing a recorded sound. In both cases, what the listener is hearing are sound waves reaching his or her ears.

difficult to achieve and why we can't really block sounds from entering our ears. Take, for example, a string telephone, a children's toy made by connecting the bottoms of two paper cups with a tautly held string. A simple device like the string telephone makes it possible to have a conversation with someone over distances of up to 100 feet without shouting. Because the cups and the string are solid media, sound travels more effectively through them than it would through air. As a result, the users can communicate across large distances at a volume that would be inaudible if spoken through air. The efficiency with which solid media can transmit sounds is interesting from a scientific perspective, but from a philosophical and phenomenological standpoint, what is remarkable is the complete auditory transparency of the cups and the string. When using the string phone, what you hear is the voice of the person speaking into the cup. The cups and the string are in themselves perfectly inaudible.

4. Transmitting vs. reproducing sounds

When I open my window and hear the rain falling, I know I will need to take an umbrella when going out. Like the other sense modalities, hearing gives us direct access to our immediate environment, enabling us to explore and experience the world around us. Philosophers disagree about whether the falling rain is directly heard or only indirectly perceived, but most of them concur that audition plays a significant role in the way we behave in and acquire beliefs about our environment.

Things get more complex when considering acoustic devices like telephones or phonographs. Although ordinary people usually believe that they do hear the sweet voices of their loved ones over the telephone, philosophers often express some resistance to this naïve view. Because the voices heard over the telephone are mediated by the telephone's loudspeaker, philosophers tend to think that the voice heard over the telephone is not the voice of their

interlocutor, but qualitatively similar sounds produced by the telephone speaker. This view is expressed by O'Callaghan:

> Televisual and telephonic experiences thus are on a par concerning whether some form of perception takes place. In what sense, if any, do both kinds of cases involve a variety of genuine perception?
> The first apparent obstacle is that both involve at best *mediated* awareness. If you hear the sound of your conversant speaking during a telephone conversation, you hear it by hearing the sound of the telephone's speaker. (O'Callaghan 2007: 152)

As acknowledged by O'Callaghan, if the sounds heard over the telephone are not identical to the sounds produced by your interlocutor's vocal cords, one can doubt that you genuinely hear his/her voice. In fact, the belief that a phone call involves mediated access to the voice of your conversant seems inevitably to lead to the conclusion that conversing over the telephone involves a perceptual illusion.

I believe that this surprising conclusion arises mainly because the function of loudspeakers has been largely misunderstood. It is commonly assumed that they *reproduce* the sounds emitted by a distant source. But this view is highly problematic for different reasons. First, this view implies that a thin membrane or a flexible panel in a box can reproduce all existing sounds: the buzz of a bumblebee as well as the powerful and rich sounds emitted by a philharmonic orchestra of more than one hundred musicians playing different scores with different instruments. But, as clearly stressed by Kulvicki, the acoustic properties of a speaker are very different from the properties of the objects it is supposed to reproduce:

> Objects can and often do *make* sounds they do not *have*, in the sense that they can be stimulated to vibrate in a way distinct from how they are disposed to vibrate in response to being thwacked. This leads to illusory impressions of objects' sounds brought on by the uncharacteristic pressure waves they produce. The most obvious examples of this phenomenon are loudspeakers. What does a speaker sound like? If you are asking about the speaker's sounds, the right answer is "Not much." Test this by thwacking your speakers (be

> gentle!). At best, you get a dull, short-lived thud. Speakers respond equally poorly to stimulation across the range of audible frequencies. This fact about speakers makes it easy to drive them to make the sounds you want them to make. (Kulvicki, 2008: 6)

Contrary to Kulvicki's remark, it doesn't seem that the physical characteristics of loudspeakers are selected for their ability to *reproduce* sounds but rather for their capacity to *transmit* sounds. Unlike sounding objects, the frequency at which speakers push air back and forth is caused by an electrical signal and does not depend upon the material properties of the speaker itself. This is why the shape and the material of the diaphragm of the speaker are selected not for their acoustic properties but for their mechanical properties. The diaphragm of the speaker is indeed constructed of a thin membrane that combines rigidity with lightness of weight in order to push air back and forth as rapidly as needed.

This point is crucial, because it explains why speakers are perceptual media and not, as habitually supposed, a special source of sounds. As stressed by Heider, perceptual objects have a unity that perceptual media lack. To explain how the medium contributes to perception without interfering with the information it conveys, Heider distinguishes between perceptual objects, which are internally conditioned, and perceptual media, which are externally conditioned. The fact that media are externally conditioned corresponds to the fact that their parts are causally independent of each other. Any air molecule can, for example, move freely without affecting the way the other air molecules behave. By contrast, all the parts of an internally conditioned object are interdependent. As Heider (1959:4) writes:

> The process on the surface of the stone, which reflects the light rays, is a process which is conditioned by the substratum . . . the fact that this particular kind of process occurs, namely, one which contains waves of particular lengths arranged in certain patterns, is determined by properties of the stone. The process in the medium, on the other hand, is conditioned externally. What happens in it is dependent on the form of the impinging process; the special state of the medium is to a high degree irrelevant for the form of the process in it.

According to Heider, the causal independence of the medium's constitutive parts explains how information can be transmitted from the perceptual object through the medium without interference. Because the medium's parts are causally independent of each other, the medium as a whole can remain undisturbed by a particular process even while the medium's parts are directly affected by it. He clarifies (1959:5) this idea in the following way:

> Vibrations whose form is conditioned externally are called “forced vibrations”, those whose form is conditioned internally are called “free vibrations.” Forced vibrations have besides the fact that they are conditioned externally, another characteristic, one which they have in common with most externally conditioned processes. Let us consider two cases of movements of a ball.
>
> In one case a ball is pushed so that is rolls across a plane. In another case, the ball is guided by hand and its movements are dependent at moment on the movement of the hand. We are certainly justified in saying that the movement of the ball is more externally conditioned in the second case. (…)
>
> In one case, the event forms a unit, that is, one part causes the next and is caused by the previous one. The parts of the movement are dependent on each other and cannot be separated from each other. In the second case, on the other hand, the event is a composite one, the single parts are to a high degree independent of each other, and there is no causal connection between the parts of the event since each part is caused separately from the outside.
>
> All forced vibrations are such composite events in which a continuous influence is exerted from the outside. The vibration is guided by the external cause in each small section.

Heider’s example of the hand guiding a ball illustrates perfectly the functioning of loudspeakers. The vibrations of a speaker are not internally conditioned, as are those of a vibrating tuning fork. The movements of the diaphragm of a speaker are caused by electrical signals that *constantly* control the speed and amplitude of the pushing-and-pulling motion of the diaphragm. As with Heider’s guided ball, the movements of a speaker are dependent at all moments on an outside cause. The resulting loudspeaker’s complex movement is therefore composite, in the sense that each

part of this complex movement is independent of the others, and does not exhibit a unitary structure.

Heider's approach to perceptual media is significant for the philosophy of perception in that it provides a physicalist criterion with which to distinguish between perceptual objects and perceptual media. But its importance is even more obvious when considering how the notion of perceptual media can explain how the medium carries information about the perceptual object to the perceiver. According to Heider, the independence of the medium's parts explains why certain materials can be informational vectors. He explains (1959:7):

> Now we can understand why certain processes can serve as mediators. The light rays which meet the eyes are messengers from the object and represent it. . . . The mediator processes which meet our sense organs are spurious units; they have unitary form not because of their own character but because they are coordinated to objects. . . . [The light rays] contain a strict order which cannot be attributed to the waves themselves since they are independent of each other. The situation is very similar in the case of the order in time of sound waves.

If air, water, wood, and speakers can serve as auditory media, it is because they are all capable of preserving and transmitting the specific perturbances caused by an initial vibrating object. In fact, the characteristic way in which the auditory medium's particles vibrate depend directly on the material and on the way the source is stimulated. This is the reason why the audible qualities we hear, such as pitch and timbre, are correlated to physical features of sound sources and not to some physical characteristic of the auditory medium. Although the physical properties of the medium can affect audition by selecting the sound waves that can reach the ears of the listener, they do not fundamentally determine the audible qualities of sounds. Bricks, for instance, absorb or reflect high frequencies but have limited absorption at low frequencies. This fact can explain why a wall can protect you from frequencies corresponding to human voices but can be ineffective in the case of low frequency noise. The physical characteristics of the medium, however, do not affect the characteristic features of the sounds you can hear through a wall.

I have argued that loudspeakers do not to reproduce sounds but rather transmit sounds like any other auditory medium. This new approach to the role of speakers has different interesting outcomes. The first interesting consequence is that viewing speakers as auditory media reconciles the commonsense view of telecommunication devices, like the telephone or radio, with a philosophical and scientific view of sounds and acoustic properties. By understanding speakers as auditory media rather than sound replicators, one can justify the commonsense belief according to which the sounds we hear over the phone are identical with the voice of the person on the line and not merely a technological simulacrum.

Another significant advantage of the view defended here regarding the nature of loudspeakers is that it dissolves the problem regarding sound reproduction. If speakers reproduce distant sounds, sounds emitted by radios and telephones are numerically different from the sounds they reproduce. The problem is therefore to explain how the sounds emitted by radios and telephones can reproduce sounds very different in nature. If sounds emitted by radios and telephones are different from the distant sounds they reproduce, the difficulty is to explain how the former relay information about the latter.

To explain how sounds produced by loudspeakers can relay information about other sounds, it is commonly suggested that the speakers' function is to imitate other sounds. Kulvicki, for example, writes:

> An audiophile's recording of the event might, upon playback, imitate the sound so well as to fool blindfolded listeners. This trompe l'oreille reproduces the sound, on this view, and this explains why people are fooled. (2017: 91)

The problem is that there is no evidence supporting the view that speakers can imitate or reproduce any sounds other than the sounds they would produce when hit with an object. Suppose you are listening to Martin Luther King Jr.'s speech delivered on August 28, 1963, on the radio. According to the imitation view of sound reproduction, the sounds produced by the radio on this

occasion are similar to the sounds of Martin Luther King Jr.'s voice. However, the physical properties of a radio are so different from those of the human vocal tract that it is impossible for a radio to reproduce the sound of a human voice. Consider, by way of contrast, the ability of certain birds to mimic human speech. If parrots and other birds can mimic human speech, it is because the vocal organ of birds, the syrinx, although very different from the human larynx, uses the same physical mechanism to produce sound. [6] Headphones and loudspeakers, on the contrary, do not exhibit any comparable mechanism: they do not display, for example, an airstream mechanism like the pulmonary system or any self-oscillation apparatus like the vocal cords or the walls of the syrinx.

Because the sounds we hear when listening to the radio come from the radio's speakers, philosophers have concluded that those sounds are produced by the radio, while they are in fact just transmitted by it. When moving without obstruction in a medium like air, soundwaves propagate unidirectionally from the source to the listener. It is therefore natural for us to exploit our ability to perceive the direction of sounds to infer their location. But like any wave, a soundwave doesn't stop when it reaches the end of a medium or encounters an obstacle. In such cases, its path can bent through reflection, diffraction, or refraction. For example, when travelling through a tube, the soundwave's trajectory depends on the shape of the tube. It is, for example, impossible to localize the sounds perceived through a stethoscope unless one can localize the chestpiece at the end of the stethoscope. As we will see in §5, when philosophers rely on the direction of the sound coming from a loudspeaker to localize its source, they simply omit situations in which sounds do not travel in a straight path and falsely conclude that they hear the sound of headphones or a loudspeaker.

[6] Elemans, C.P.H. & al. (2015)

5. The nature of recorded sounds

Considering speakers and headphones as auditory media rather than sound replicators justifies our commonsense beliefs about the veridicality of our auditory experiences when using phones and radios. Perceptual instruments, like mirrors, telescopes, microscopes, etc., have extended our perception beyond its natural limits, and telecommunication and audio devices are no exception. Like the telescope, which has extended our vision far beyond the horizon, the radio and telephone have enabled us to perceive sounds across oceans and continents and even through the space. Thanks to these new tools, the spatial limits of our perceptions are constantly expanding, but so are their temporal limits. I will argue indeed that when we hear recorded sounds, like the "I have a dream" speech delivered by Martin Luther King Jr. in 1963, we directly hear sounds of the past. We directly hear, for example, exactly the same sounds we would have heard if we were standing near the Lincoln Memorial on August 28, 1963. Although this view of the nature of recorded sounds is in accordance with the phenomenology of these experiences, most philosophers have rejected this straightforward view of audio recording. They have argued instead that these experiences are essentially different from standard auditory experiences and that they cannot be considered genuinely veridical. O'Callaghan maintains, for instance, that auditory experiences fall on a "veridicality" continuum that starts with unassisted hearing of actual and nearby sounds, which is genuinely perceptual, and ends with the hearing of recorded sounds, which cannot be considered "fully" perceptual because such sounds are impoverished. He explains:

> According to my "illusory perspectives" account, hearing the recording in your living room is a way of perceiving the sound of Hendrix performing "The Star-Spangled Banner." Your experience of that sound, however, is impoverished and illusory in a number of respects, including its spatial and temporal perspectival contents. (O'Callaghan 2007: 160)

Although I agree with O'Callaghan's claim that audition mediated by audio devices can be impoverished in comparison to auditory experiences enjoyed in natural settings and that the spatial

and temporal perspectives provided by recorded sounds are singular and sometimes misleading, I do not think that listening to recorded sounds is illusory in any sense.

"Normal" audition usually takes place in an open space where sound localization is not strongly affected by absorption, reflection, diffraction, or reverberation. But there are a number of things that can obstruct our capacity to correctly localize sounds. For example, although the reverberation in a room can increase the audibility of sounds, and therefore enhance the listening experience, it often reduces the listener's ability to correctly localize the source sounds. In acoustics, like elsewhere, there is no free lunch. As rightly noted by O'Callaghan, the use of audio devices dramatically reduces our ability to locate the source of sounds. The first reason is somewhat trivial. Audio devices are generally used when certain sounds are not audible in standard conditions. It is not possible to point to or gaze in the direction of those sounds, simply because they are not situated in our vicinity. If one takes seriously the proposal that we hear sounds transmitted through the audio device and not merely their replica, it is quite understandable that we cannot locate their source in our immediate environment, since they are beyond our "natural" auditory reach. Consider, for instance, the speaking pipes that were used in 19th-century houses and offices to communicate between distant rooms. Although those pipes would enable listeners to hear their conversant from a distance of up to 950 m, there would be no way for them to know just by listening to their conversant's voice where the latter was located. As with the use of speaking pipes, the sounds transmitted by audio devices provide insufficient cues to enable the listener to locate the source. O'Callaghan is therefore right to claim that the spatial properties perceived through mediated hearing are poorer than those accessible through direct hearing. Unlike our capacity to localize a bird by its singing, a voice heard over the phone provides no clue about the location of our conversant. O'Callaghan is wrong, however, to claim that the lack of spatial information in the case of mediated hearing involves an illusion of spatial perspectives. According to O'Callaghan,

recorded sounds are illusory because they mislead the hearing subject about where and when the sounds are located. He says:

> Hearing a good sound recording presents information about the locations of the musicians relative to the microphone. Though experienced perspectival content varies with changes to the perspectival relationship between the camera or microphone and the original object, event, or sound, it misrepresents or fails to reveal your actual perspectival relationship to the source. Our resistance to accepting such experiences at face value is learned to the extent that we have learned to distinguish the mirror's, camera's, or microphone's perspective from our own. However compelling, the experience in real time of a transmitted sound or image is perspectivally illusory because the subject does not in fact occupy the point of view of the microphone or camera. (O'Callaghan 2007: 155)

We must notice first that O'Callaghan's remark supposes that localizing sounds coming out of speakers relies on the same capacities as those in use when hearing sound in a "natural" setting. In other words, O'Callaghan supposes that the listener is always epistemically innocent.[7]. But is that really so? As stressed by Prinz (2012: 177) regarding vision with a mirror, the perceptual experience would be illusory only if the subjects were inevitably fooled by their use of mirrors. But this is not the case. When mirrors are correctly used, perceivers do not experience objects in mirrors as mislocated. He explains:

> The 'seeming' here is not visual. The world may look inverted, but relearning motor skills makes it possible to behave as if things had their standard orientation. Compare what happens when you become adept at combing your hair in a mirror. You learn to move your hand backward to reach the back of your head even though the mirror reflection suggests that you should move your arm forward, since the reflected back of your head is in front of you. When you master this skill, the mirror doesn't appear inverted.

[7] See Casati 2012

The difference between innocent and un-innocent mirror perception is based on the way vision is coordinated with the motor system and the other sense modalities. Because perception and motor action are closely related, any change in this perceptual-motor coordination requires some adaptation. Adjusting our behavior to new visual information is required in all situations in which a new perceptual medium is introduced. Such situations include those in which people use mirrors, but also those in which they use magnifying or shrinking lenses and displacing or rotating prisms. All of these situations differ from those associated with "normal" perception, according to the ways in which the plurality of frameworks associated with the different sense modalities and motor systems are coordinated. The fact that an adaptation is required for a subject to correctly localize the objects that he or she perceives with a mirror does not mean that his or her visual experience is illusory or erroneous. What it shows is that visual and motor frameworks need to be aligned: pointing to an object that you perceive in open air and pointing to an object that you perceive using mirrors or prismatic lenses rely on different forms of visual-motor coordination. But fortunately, most perceivers adapt their behavior to the changes of perspectives caused by mirrors.

Similarly, there is no illusory spatial perspective when we listen to Glenn Gould's recording of the *Goldberg Variations*. We are not surprised to find a loudspeaker in the room instead of a piano. As nicely put by Casati about the idea that mirrors generate systematic illusions : "It is not content that is illusory; the illusion is the impression of having illusory content."

Although I have argued that recorded sounds do not generate an illusory perspective, they are certainly capable of producing errors and false judgments. If I hear a record being played in my neighbor's apartment, I can falsely believe that there is a piano behind the wall. But experiences with recorded sounds can be even more vivid and disconcerting. Binaural recording, for example, which uses two microphones on a mannequin's head to recreate the human auditory system, can be

so realistic that it is difficult not to believe that you are hearing sounds around you.[8,9] If hearing recorded sounds is not illusory, how can we account for the misleading experiences associated with them? Why should we resist Kulvicki's conclusion that recorded sounds are kinds of "trompe l'oreille"?

One reason to resist the idea that perceiving recorded sounds is illusory is to avoid having to introduce any intermediaries to explain the phenomenal character of these experiences. According to the approach proposed here, distant or past sounds are directly accessible in our auditory experiences of recorded sounds. We not only have the impression of hearing Martin Luther King Jr.'s voice when hearing a recording of his speech, we actually hear his proper voice. This view fits nicely into the most general framework provided by naïve realism. According to naïve realism, there is no need to introduce representations, sense data, or any mind-dependent entity to explain the phenomenal character of perceptual experiences. According to this view, what we perceive is a fragment of the world itself. Reference to mind-independent objects and properties is all that is needed to account for the phenomenal character of perceptual experiences. The characterization of recorded sounds offered in this paper defends the same model. It explains how auditory media modify our perceptual experiences without introducing mind-dependent or subjective entities. It also explains how they can enrich our auditory world and extend our auditory space.

Although naïve realism has many phenomenological and epistemological merits, it also seems to face some serious difficulties. One of them is the possibility of illusions and perceptual errors. In effect, if perceptual experiences are constituted only by mind-independent objects and properties, how can there be illusory or erroneous experiences? How can perceptual experiences be constituted by things that are different from what they really are? Naïve realism seems inadequate

[8] For some nice examples of binaural recordings, see https://binauralenthusiast.com/examples/.
[9] Surprisingly, binaural recording is not a recent invention. In fact, in 1881, Clément Adler demonstrated a transmission system that allowed listeners to enjoy the opera binaurally by holding one receiver against each ear. This broadcasting system was later used in several European cities for transmitting opera performances and plays.

to explain illusion and misperception, because it does not admit a dichotomy between what things are and how they appear.

Like many nonstandard media, communication tools and audio devices have been understood as generating illusions and misperceptions.[10] If naïve realism cannot distinguish between veridical and nonveridical perceptual experiences, how can it deal with these apparent cases of illusions? The strategy I have proposed here is to contest the claim that audio devices generate perceptual illusions. I have suggested that apparently illusory or erroneous perceptual experiences can be accounted for in terms of past and distant sounds. What is missing, however, is a plausible account of their misleading character. Misleading appearances have typically been understood as involving illusory or erroneous perceptual experiences, but there are other options. In particular, it is possible to account for misleading appearances in doxastic terms. This view has been defended by Arthadeva (1960) and, more recently, by Genone (2014) and Mizrahi (2019).

Perceptual judgments rely on what is accessible through particular perceptual experiences but also on a rich contextual and sensorimotor knowledge. For example, determining whether an object is stationary or moving relies on what is perceived as well as on the perceiver's implicit knowledge of his own movements. The predominant role of background knowledge in perceptual judgments is particularly salient in specular perception. As recognized by Casati, the perceiver's awareness of the fact that he or she is dealing with a mirror removes the misleading character associated with mirror perception:

> Looking in a mirror is not (normally) like looking through glass, rather it is obtaining information about part of the world that sends to the mirror the light that the observer uses to perceive the scene. Captain Hook is not seen erroneously on the other side of the wall on which the mirror hangs, but on the same side of the room where the observer is located; and he is seen where he is precisely because it is known that a mirror is in use. (Casati 2012:201)

[10] To take only a few examples: microscopes make germs appear bigger, telescopes make planets seem closer, mirrors make hands appear reversed, etc.

Knowing whether or not we are dealing with a mirror seems to determine whether or not we can accurately identify the location of the objects we perceive. The absence of this piece of knowledge therefore seems to be essential to the misleading character associated with specular perception: once we know we are dealing with mirrors, we are not inclined to believe that an object is located behind the mirror or that our body is right–left reversed.

The case of recorded sounds is very similar, because recorded sounds introduce systematic changes in the way auditory experiences are related to other sense modalities and to the motor system. In natural settings, the visual and auditory systems work together to identify and localize objects and events. When we knowingly hear recorded sounds, however, our dispositions change and an adaptation is therefore required for the subject to align his or her auditory experiences with the information coming from the other senses and to engage in appropriate behavior. As clearly identified by Chion (1983: 18), modern technology involves a different way of listening to sounds. He explains:

> *Acousmatic*: a rare word, derived from the Greek, and defined in the dictionary as: *adjective, indicating a sound that one hears without seeing what causes it.*
> The word was taken up again by Pierre Schaeffer and Jérôme Peignot to describe an experience which is very common today but whose consequences are more or less unrecognized, consisting of hearing sounds with no visible cause on the radio, records, telephone, tape recorder etc.
> Acousmatic listening is the opposite of *direct* listening, which is the "natural" situation where sound sources are present and visible.
> The acousmatic situation changes the way we hear. By isolating the sound from the "audiovisual complex" to which it initially belonged, it creates favorable conditions for *reduced listening* which concentrates on the sound for its own sake, as sound object, independently of its causes or its meaning (although reduced listening can also take place, but with greater difficulty, in a direct listening situation).

Perception is commonly accompanied by some control-oriented monitoring that grounds our knowledge about our own perception. So, when confronted with auditory experiences that differ from ordinary perceptions, the control mechanism, whose function is to monitor perceptual processes, can either make us aware of the changes produced by new perceptual media or fail to notice those changes. In the first case, the perceptual experiences will not be prone to generate false belief. In the latter case, they will. Imagine a scenario in which a subject who was not aware that he or she was hearing a recorded sound suddenly understands that the sound is coming out of a loudspeaker. How does this knowledge affect the listener's experience?

I suggest that the difference experienced before and after noticing the presence of a loudspeaker is not auditory; it is explained in terms of the practical skills and cognitive processes exploited by the subject in localizing objects in his or her surroundings. Although the objects are qualitatively identical at the auditory level, people can misjudge their location, because the various sensorimotor mechanisms that ground the way they localize objects in their environment have been altered. Contrary to the prevailing account of recorded sounds, I therefore suggest that mislocations related to recorded sound are not perceptual but doxastic. Rather than adapting his or her auditory experience to an acoustamic situation, a perceiver can misleadingly rely on the direction of a sound and draw a false inference about its location.

The same response can be made to the idea that recorded sounds generate temporal illusions. It is possible to draw false inferences when listening to recorded sounds. And falsely believe that what you are hearing takes place at the same time you are experiencing it. But most of our perceptions of recorded sounds do not appear this way. Like photographs,[11] recorded sounds can send us back in time, eliciting memories and nostalgia. Hearing sounds of the past through recorded sounds can be intense and emotional, but there is no temporal illusion involved. If recorded sounds can indeed evoke a longing for days long gone, it's because we know that those sounds belong to

[11] For similarities between photography and recorded sounds, see §6.

the past. The great gift of Scott de Martinville and Edison is not to have invented devices able to reproduce vanished sounds but rather to have created technological instruments literally able to store and transmit them to future listeners.

6. Conclusion

Although audiophiles[12] would probably claim that some hi-fi systems deliver a better sound than anything that can be experienced in a concert hall or a jazz club, many purists will disagree and maintain that natural acoustic sounds experienced in a concert hall cannot be surpassed. The truth is that the choice of a perceptual medium is never right or wrong simpliciter; it is only appropriate or inappropriate to particular goals. Perceptual media are crucial components of perception, and their choice has a critical impact on what is perceived. For example, the microscope, which incorporates a magnifying lens, is perfectly suited to observing chromosomes or bacteria, but very ill suited to spotting birds in a garden. Tradeoffs are always involved when choosing auditory media, because it all depends on what sounds or what auditory qualities we want to perceive. For example, concert halls that are designed to make a single soprano sound as loud as a full orchestra are not suited to rock concerts. And auditoriums used for music are not always appropriate for conferences; that is why acoustic enhancement is sometimes necessary. All these examples show that there is no perfect auditory medium but only media tailored to different purposes.

Technology has extended the diversity of auditory media and opened new routes for sounds to reach our ears. In contrast to the idea that auditory experiences enabled by new technological means are illusory or inferior, I have claimed that hearing through prosthetic or audio devices is not ontologically different from natural hearing. In fact, it is possible to argue, like Massin (2010: 103), that there is a continuum between the perceptual system and the perceptual media located outside the perceiver's body:

[12] An audiophile is a person who is enthusiastic about high-fidelity sound reproduction.

> According to present suggestion, the concept of medium can be extended to some of the perceiver's body parts, in particular to his perceptual system. If air or eyeglasses belong to the medium, why is it not the same for the cornea or the retina? Why not include also the optic nerve, and the primary visual area in the causal medium which keeps us and the object apart? Is it not somewhat arbitrary to consider that the causal medium ends as soon as the causal flux enters the body?

As argued in this paper, perceptual media affect our perception by selecting what portion of reality is perceived. The choice of a medium is not therefore right or wrong simpliciter; it is only appropriate or inappropriate to particular perceptions. The human auditory system, which incorporates different auditory media (the eardrum, the ossicles, endolymph,…) is perfectly suited to hearing a singing bird within hearing distance, but it fails to perceive sounds of the past or sounds located across vast distances.

By recording sounds, we make those sounds accessible to future and distant listeners, in the same way photographs preserve the look of objects for future and distant viewers. The view of recorded sounds presented here does indeed bear some very strong resemblances to Bazin and Walton's view of photography:

> Amid Bazin's assorted declarations about photography is a comparison of the cinema to mirrors. This points in the right direction. Mirrors are aids to vision, allowing us to see things in circumstances in which we would not otherwise be able to; with their help we can see around corners. Telescopes and microscopes extend our visual powers in other ways, enabling us to see things that are too far away or too small to be seen with the naked eye. Photography is an aid to vision also, and an especially versatile one. With the assistance of the camera, we can see not only around corners and what is distant or small; we can also see into the past. We see long deceased ancestors when we look at dusty snapshots of them. To view a screening of Frederic Wiseman's *Titicut Follies* (1967) in San Francisco in 1984 is to watch events which occurred in 1967 at the Bridgewater State Hospital for the Criminally Insane. Photographs are *transparent*. We see the world *through* them. (Walton 1984: 271)

I suggest that the similarities between my view and Bazin and Walton's represent more than a mere coincidence: they indicate that the approach to recorded sounds defended in this paper should be extended to all technologies that enrich our perceptual world. More than just supports for our imagination or modes of representation, these new technologies extend our perceptual capacities by expanding our perceptual fields. The taste of a madeleine can bring back childhood remembrances, but photographs and recorded sounds can do more than just trigger vivid memories. They enable us to relive parts of our childhood by providing everlasting access to our past and capturing thereby the ephemeral for eternity.

7. References

Arthadeva, M. 1960. "Mirror images" are physical objects: A reply to Mr. Armstrong, *Australasian Journal of Philosophy*, 38:2, 160-162.

Chion, M. 1983. *Guide des objets sonores : Pierre Schaeffer et la recherche musicale*, Paris : Buchet/Chastel.

Casati, R. 2012. Illusions and Epistemic Innocence, in C. Calabi (ed.), *Perceptual Illusions. Philosophical and Psychological Essays* (Palgrave MacMillan)

Casati, R. & Dokic, J. 1994. *La philosophie du son*, Nîmes: Chambon.

Casati, R., and J. Dokic. 2010. Sounds. In The Stanford encyclopedia of philosophy, ed. E.N. Zalta. http://plato.stanford.edu/entries/sounds/

Elemans, C.P.H. & al. 2015. "Universal mechanisms of sound production and control in birds and mammals", *Nature Communications*, 6: 8978.

Genone, James 2014. Appearance and Illusion, *Mind* 123(490), 339–76

Heider, Fritz 1959. Thing and Medium, *Psychological Issues* I:1-34.

Kalderon 2015 *Form without Matter, Empedocles and Aristotle on Color Perception*.Oxford: Oxford University Press.

Kalderon 2018 "Aristotle on Transparency". In *Perceptual Ephemera*, ed. T. Crowther and C. Mac Cumhaill, Oxford: Oxford University Press: 219-237.

Kulvicki, John, 2008. "The Nature of Noise", *Imprint* 8 (11)

Leddington, J. 2014. "What we hear". In *Consciousness Inside and out: Phenomenology, Neuroscience, and the Nature of Experience*, ed. R. Brown, 321–34. Dordrecht: Springer.

Leddington, J, 2018. "Sounds fully simplified », Analysis, pp. 1–9

Massin, O. 2010. *L'objectivité du toucher*. (Doctoral dissertation). Retrieved from http:// www.academia.edu/3010474/Lobjectivite_du_toucher

Mizrahi, V. 2019. "Mirrors and Misleading Appearances", *Australasian Journal of Philosophy*, 97:2, 354-367.

Nudds, M. 2010. “What sounds are”. In *Oxford Studies in Metaphysics*, vol. 5, ed. D. Zimmerman, 279–302. Oxford: Oxford University Press.

O'Callaghan, 2007. *Sounds : A Philosophical Theory*, Oxford: Clarendon Press.

Pasnau, R. 1999. “What is Sound?” *Philosophical Quarterly*, 49: 309–24.

Prinz, Jesse J. 2014, *The Conscious Brain: How Attention Engenders Experience*, Oxford: University Press.

Smith 2002. *The Problem of Perception*, Cambridge, MA: Harvard University Press.

Walton, K. 1984. “Transparent Pictures: On the Nature of Photographic Realism”, *Critical Inquiry*, Vol. 11, No. 2, pp. 246-277.