

THE NATURE OF TIMBRE

VIVIAN MIZRAHI

University of Geneva

Along with pitch and loudness, timbre is commonly described as an audible property of sounds. This paper puts forward an alternative view—that timbres are properties of auditory media. This approach has many advantages. First, it accounts for the frequent attribution of timbres to objects that do not have characteristic sounds. Second, it explains why timbres are attributed not only to ordinary objects, like musical instruments, but also to surrounding spaces and architectural structures. And finally, it provides an original solution to the timbre-constancy problem.

1. Introduction

As the recent literature demonstrates, the topic of sounds, neglected for a long time by philosophers, is today at the center of rich and vivid debates in the philosophy of perception. The renewed interest of philosophers in sounds and auditory perception has fueled extensive discussions regarding the nature of sounds and their phenomenology. Moreover, the original ontological theories recently proposed provide a diverse palette of views regarding the relations between sounds, events, objects, space, and time as well as stimulating new explanations for the specificity of the auditory modality. Philosophers disagree about the nature of sounds for many different reasons. They debate, for example, about whether sounds are properties (Pasnau 1999; Kulvicki 2008; 2014; Leddington 2014; 2019), particulars (Casati & Dokic 1994; 2010; O'Callaghan 2007), or abstract individuals (Nudds 2009b), but also discuss their location relative to the hearer by confronting proximal, medial, distal, and aspatial theories of sounds.

Although these numerous recent developments in the philosophy of sound have undeniably contributed to a better grasp of the challenges posed by sounds and their spatial properties, it is remarkable that very little progress has been made in the understanding of their qualitative dimensions, which many music

Contact: Vivian Mizrahi <vivian@mizrahi.ch>

specialists and auditory scientists consider to be at the heart of our auditory experiences.

Although sounds can differ in their apparent spatial and temporal localization and in their respective durations, it would not be an exaggeration to say that it is audible qualities that fundamentally distinguish one sound from another. Consider Dickens's description of Smithfield's meat market:

The whistling of drovers, the barking dogs, the bellowing and plunging of the oxen, the bleating of sheep, the grunting and squeaking of pigs, the cries of hawkers, the shouts, oaths, and quarrelling on all sides; the ringing of bells and roar of voices, that issued from every public-house; the crowding, pushing, driving, beating, whooping and yelling; the hideous and discordant din that resounded from every corner of the market. (Charles Dickens, *Oliver Twist*, 1838)

Dickens portrays the market as a place filled with cacophony and great confusion, but the specificity of each sound or family of sounds composing the auditory environment he describes is easy to imagine. Whistlings, barkings, bleatings, squeakings, cries, and ringings represent quite distinct experiences that are not easily confused with one another. Yet the particularity of such auditory experiences is not grounded in the intensities of the sounds heard or the way they are represented in space but rather in their characteristic audible qualities.

"Timbre" has been used to refer to what have been taken to be qualities of sounds since the mid-eighteenth century. Usually described with the help of metaphors based on other sensory modalities, the term is often used to contrast the sounds produced by different musical instruments. This is the case, for example, in Rousseau's entry for the *Encyclopedia* of Diderot and d'Alembert:

A sound's timbre describes its harshness or softness, its dullness or brightness. Soft sounds, like those of a flute, ordinarily have little harshness; bright sounds are often harsh, like those of the *vielle* or the oboe. There are even instruments, such as the harpsichord, which are both dull and harsh at the same time; this is the timbre. The beautiful timbre is that which combines softness with brightness of sound; the violin is an example.

Although descriptions of timbre that, like Rousseau's, use visual and tactile metaphors are still in use, the American National Standards Institute (ANSI) adopts Helmholtz's approach and gives a negative definition of timbre. Allegedly more rigorous than impressionistic, positive verbal analogies, the ANSI definition of timbre is "that attribute of auditory sensation in terms of which a listener can

judge that two sounds similarly presented and having the same loudness and pitch are dissimilar” (ANSI 1960: 45, §12.9).

This characterization of timbre has often been criticized (Bregman 1994; Sankiewicz & Budzynski 2007; Donnadieu 2007) and is defective in many ways. As Bregman explains, this definition is both too specific and too general. It is too specific because it leaves out the timbre of sounds that lack pitch and the timbre of sounds whose pitch cannot vary:

This is, of course, no definition at all. For example, it implies that there are some sounds for which we cannot decide whether they possess the quality of timbre or not. In order for the definition to apply, two sounds need to be able to be presented at the same pitch, but there are some sounds, such as the scraping of a shovel in a pile of gravel, that have no pitch at all. We obviously have a problem: Either we must assert that only sounds with pitch can have timbre, meaning that we cannot discuss the timbre of a tambourine or of the musical sounds of many African cultures, or there is something terribly wrong with the definition. (Bregman 1994: 92)

But the definition is also too general, because it cannot distinguish between what seems to genuinely characterize the audible qualities of a sound and what seems to be presented only accidentally when we hear a specific timbre. Bregman explains:

We can judge the dissimilarity of a sound presented to the right ear from one presented to the left and yet any fool knows that this is not a difference in timbre, so we can conclude that they were not similarly presented. Does a tone pulsed twice per second differ in timbre from one pulsed three times per second? Probably not; therefore since they are dissimilar in some way, they must not have been “similarly presented”. This argument implies that rate of presentation must be held constant. But if we accepted that conclusion, we could never discover the obvious truth that tones pulsed 20 times per second are different in timbre from those pulsed 40 times. I think the definition of timbre by the American Standards Association should be this: “We do not know how to define timbre, but it is not loudness and it is not pitch.” (Bregman 1994: 92–93)

According to Bregman, the notion of timbre as defined by the ANSI is an “ill-defined wastebasket category” designed to include in bulk every audible quality except loudness and pitch. He proposes therefore to drop this notion until “a better vocabulary concerning timbre” is developed. Bregman is probably right to stress that a better definition of timbre is needed. But such a definition cannot

be furnished without a prior understanding of the nature of timbre and of the audible characteristics to which it applies.

The nature of timbre would not be so problematic if timbre were correlated only with audible qualities independent of loudness and pitch. What makes timbre particularly difficult to characterize is the fact that it is also a way to classify sounds in a positive way. Consider Rousseau's definition and his paradigmatic use of the notion of timbre to distinguish sounds of different musical instruments or families of instruments. Musical instruments, unlike most sounding objects or materials, can produce steady sounds at different pitches.^{1, 2} Unlike a tuning fork, for example, which resonates only at a specific frequency, musical instruments have been designed to produce steady sounds in a large range of frequencies. This is why, for example, oboes and violins can play the same line of music in unison. Now, as stressed by Rousseau and the ANSI definition, the fact that listeners can tell apart the sounds of violins and oboes regardless of their pitch and loudness is certainly notable, but what is remarkable is the fact that despite differences in their pitch and loudness, certain sounds appear to have some obvious similarities. Irrespective of whether a clarinetist plays an A, B, or C, there is a distinctive audible quality shared by the clarinet's sounds, that is, a quality that differs from that of sounds produced, say, by a violin.

Now, timbral similarities are granular; that is, similarities between timbres apply at different levels. Sounds produced by string instruments, for example, can exhibit timbral similarities that make it easy to distinguish them from sounds produced by wind instruments. These sounds are, however, dissimilar at a different level, because it is possible to discriminate among wind instruments. At this level of granularity, the timbral qualities of cellos and violins can be told apart and consequently perceived as dissimilar. Finally, individual instruments exhibit timbral characteristics that allow the listener to discriminate between them and eventually to recognize their source. This is the case for top violin makers, who can craft instruments tailored to their future owners. At this bottom level, it is in theory possible to distinguish the timbral characteristics of different violins.

Although timbres are easily recognized and distinguished from each other, the questions "What are timbres?" and "What are the bearers of timbres?" remain unanswered. Philosophical approaches commonly identify timbre with an audible property of sounds. I argue in what follows that timbre is not a property of sounds, but rather a characteristic of auditory media. I maintain that this new perspective provides a better explanation of timbre constancy and accounts

1. "The steady pitch is an idealization that probably was discovered when humans invented musical instruments" (Bregman 1994: 104).

2. Notice also that many sounds do not have any specific pitch: a door slamming, rain falling or a hammer hitting a nail for example.

for the wide variety of bearers of timbre, which include not only resonant objects but also empty spaces and architectural structures.

I begin this investigation of the nature of timbre in §2 with a brief exploration of the variety of objects that have timbral properties. This inquiry reveals the close ties not only between timbres and the sources of sounds, but also between timbres and auditory media, such as empty spaces and architectural structures. In §3, spell out an Aristotelian account of auditory media and explain their role in hearing. I then consider in §4 philosophical views on audition and auditory media and explain in §5 how objects that do not have audible properties can nevertheless have timbral characteristics. I draw attention in §6 to the benefits of a theory that attributes timbre to auditory media rather than to sounds or sources of sounds and conclude the paper in §7 with a short overview of the relations between timbre and music.

2. Timbres, Sounding Objects, and Auditory Media

Along with pitch and loudness, timbre is commonly described as an audible quality of sounds. But although pitch and loudness can be and often are shared by sounds originating from very different objects, this is not typically the case with timbre. As is often stressed, timbral properties are closely associated with the sound-producing objects and their constitutive materials (Handel & Erickson 2004). Violins and oboes can play the same note, but it is usually not difficult to distinguish their respective timbres. Conversely, two different notes played on a violin are typically perceived as similar in timbre.

Timbral qualities are often described directly by reference to the objects that produce sounds with typical timbre. Even in musicology, referring to the timbral characteristics of an instrument almost always involves mentioning the instrument itself. Instead of using a specialized vocabulary to describe timbral characteristics we just rely on our shared experiences with their source and refer to timbres according to the way they are produced—“he sounds just like Elvis Presley” (Handel & Erickson 2004: 588).

The close relation between timbral properties and the physical properties of objects seems to support physicalist and objectivist views, which consider sounds to be objective elements of the world that exist independently of listeners and their experiences of hearing. For Kulvicki, who defends the idea that sounds are stable dispositions of objects to vibrate, our capacity to identify an instrument through its characteristic timbre despite variations in the individual notes played supports “a stable property view” of sounding objects. Similarly, O’Callaghan argues that the constancy of timbres across sounds and circumstances is best explained by “features of the source and the characteristic man-

ner in which it disturbs the medium” (O’Callaghan 2007: 89). Even though not all philosophers endorse objectivist views of sounds, they assume that timbres are closely related to the sources of sounds, because timbres seem ineluctably to carry information about a particular source.

An important category of timbres is related to the human voice. In his seminal work, Helmholtz introduces the notion of timbre (*Klangfarbe*) by associating human voices with the timbre of instruments: “By the quality of a tone [timbre, *Klangfarbe*] we mean that peculiarity which distinguishes the musical tone of a violin from that of a flute or that of a clarinet or that of the human voice, when all these instruments produce the same note at the same pitch” (Helmholtz 1877: 10).

Hearing a human voice is a very rich experience. It not only conveys physical information about the subject to whom the voice belongs (and even often allows for its identification), but it also often reveals the speaker’s state of mind and emotions. Variations of loudness, pitch, modulation, rhythm, etc. in an individual’s voice often directly indicate or express the subject’s emotional state, independently of the possible meaning attached to these vocal utterances. Now, although these experiences are quite common and straightforward, they are philosophically puzzling. Human voices are good emotional media because they are rich and flexible and can be controlled by subjects so that they vary along multiple dimensions, such as energy, pitch, spectral features, rhythm, etc. To express their emotions vocally, subjects rely on the extensive combinatory components of the human voice and its rich auditory palette. But to perceive a person’s sadness, anger or fear, one must first recognize that the sounds one hears belong to a unique person. How, then, can sounds that differ along many different dimensions be perceived as belonging to a single person? In other words, what is a voice, and how do we identify a voice through the variety of sounds we hear?

Consider by way of contrast the sound of a bell. Each bell has a characteristic sound that does not vary along multiple dimensions. As Kulvicki points out, the sound of a bell, which results from the disposition of the bell to vibrate in response to mechanical stimulation, is stable and unique. The sound of a bell does not vary except with respect to the variations caused by the way it is struck. The stable disposition of a bell to vibrate in a characteristic way explains, it seems, the relative invariability of its sound and its apparent dependence on the physical constitution of the sounding object.

But can this straightforward view be extended to voices? What gives voices and musical instruments their expressive power is their remarkable flexibility and the extended palette of sounds they can produce. It seems that the timbre of a voice, unlike the sound of a bell, cannot be restricted to a simple disposition to vibrate in a characteristic way. The plurality of sounds corresponding to a particular voice results from our capacity to change the vibrations of our vocal folds

as well as the size and the shape of our vocal tract. A person's voice is indeed very different from the sound of a vibrating bell. As Kalderon mischievously puts it:

The sound of a stereo speaker playing is a sound that it makes but does not have. I suspect that a person's voice is more like the sound of a stereo speaker playing than the sound that it makes when thwacked. Through a series of unfortunate events, I have first hand experience of what I sound like when thwacked. I can attest it sounds nothing like my voice. Like a stereo speaker, I produce a dull thud when thwacked. (Kalderon 2017: 90)

Although voices are straightforwardly recognizable, they reveal a duality that is not so easy to capture and describe. A voice appears as a dual phenomenon that comprises the simultaneous experiences of a constant feature—its timbre—and of some audible variations. A singer's voice—the particular timbre of the singer—is what seems to remain constant when all other audible properties vary. But what is the nature of this invariance? Is timbre, as a widely held view suggests, a dimension of sound that belongs to the same class as intensity and pitch, or does it have a different nature?

Sounds can be said to be similar and dissimilar regarding their timbral characteristics in different ways. The timbre of a Stradivarius can be perceived as different from the timbre of an ordinary violin even though they are similar in comparison to the timbre of other string instruments. But timbral characteristics are not limited to objects. The acoustic environment in which sounding objects are located directly affects what we hear, as Beranek explains:

"Timbre" is the quality of sound that distinguishes one instrument from another or one voice from another. "Tone color" describes balance between the strengths of low, middle, and high frequencies, and balance between sections of the orchestra. The acoustical environment in which the music is produced affects tone color. If the hall either amplifies or absorbs the treble sound, brittleness or a muffled quality may mar the music. If the stage enclosure or the main ceiling directs certain sounds only toward some parts of the hall and not toward others, the tone color will be affected differentially. (Beranek 2004: 31)³

A concert hall provides much more than a space for music performers and their public to gather. It directly contributes to the musical performance as an acoustic

3. Although "timbre" and "tone color" are often used as synonyms, Beranek seems to distinguish here between auditory variations that can be attributed to musical instruments and voices from those that are related to the acoustical properties of a concert halls.

extension of the musical instruments being played. As Beranek spells out in his magisterial monograph on concert halls, each musical period, style, composer, and even musician requires a particular acoustic space. For example, Bach's early fugues were intended to be played in private or ducal chapels, in contrast to his large works, like the *St. Matthew Passion*, which were created for the newly built Lutheran churches (Beranek 2004: 9). Some concert halls do suffer from serious flaws, but no auditorium can be perfect per se, because different auditoriums often exhibit incompatible properties, which result in different acoustics that are "almost always with a degree of give and take, of value lost and found" (Beranek 2004: 4). Because the many choices involved in conceiving and building a concert hall are directly relevant to the sounds that can be heard within them, it seems that no concert hall can do justice to all the sounds that could be played within it.

The way concert halls affect musical performances is fascinating in itself, but I believe its philosophical relevance is more general, because concert halls are a striking demonstration of the importance of auditory media for understanding sounds and auditory perception.

In §5 and §6, I show that perceptual media are essential to timbral properties and their variations. I argue that what gives objects their timbre is not their capacity to produce sounds, as is often assumed, but their capacity to transmit sounds in a special way and make them available to hearers. First, however, in §3 and §4, I introduce and explain the role of auditory media in audition and describe their special characteristics and their roles.

3. The Nature of Auditory Media

The notion of perceptual medium is introduced by Aristotle to explain how perception at a distance is possible. He recognizes that the remoteness of perceptual objects from the perceiver requires the presence of a causal intermediary. He argues that although the interspace between the perceived object and the observer seems empty, perception at a distance is enabled by a causal intermediary connecting the remote object to the sense organ (*De Anima* ii 7 418b13–22). Since each sense is identified by its proper object, to understand how a perceptual medium acts as a causal intermediary between the proper object and the perceiver, one must first identify the proper object of the sense. Hearing is the power to perceive sounds. An auditory medium is therefore what is required for a sound to reach a perceiver. Contrasting auditory media and objects which produce sounds (hereafter "sounding objects"), Aristotle explains that an auditory medium is not audible by itself (*De Anima* 419b19) but that it "may resonate to the sound of something else that does have sound in such a way that resonance reaches a perceiver" (Johansen 1997: 151). Unlike the scientific notion of an audi-

tory medium, which is generally used to refer to the physical elastic media, such as air or water, that propagate sound waves, the Aristotelian notion of auditory medium thus extends to all material or objects that do not make their own sound but resound “to the sound of something else” (Johansen 1997: 153).

Although a material’s properties, like its stiffness and density, influence the speed at which sound waves travel, almost any material can transmit sound and therefore serve as an auditory medium. The diversity of acoustic media explains why complete silence is extremely difficult to achieve and why we can’t really block sounds from entering our ears.⁴

The physical properties of materials are important to understand how sounds are transmitted to the perceiver, but so are the shape, the size, the constitution, and the superficial properties of the objects that surround the perceiver. To explain how a sound reaches a perceiver, one must indeed understand how sound waves are propagated through different materials, but also how they reach a particular location. Which sounds we can hear at a particular place depends indeed on the way sound waves bounce off surrounding surfaces, are absorbed or deviated in the direction of the perceiver.

Philosophers and scientists often restrict the use of the terms “auditory media” to physical elastic media, such as air or water, that propagate sound waves. Following Aristotle’s view of auditory media, I extend the notion of auditory medium to all the causal intermediaries that transmit sounds to the listener. In this wider sense, “auditory media” refers not only to the physical materials, such as air and water, that fill the space where audition takes place but also to the surfaces and objects that fill this space. I also include in its extension the hearing system itself which captures the vibratory patterns in the environment and transmit them to the brain through neural-electrical impulses, as well as to the hearing aids and technological devices, such as radio or audio instruments, used to store and transmit sounds to distant and future listeners (see Mizrahi 2020).⁵ I believe that this extended sense of “auditory media” accords better with the Aristotelian insight that perceptual media are characterized by their role as causal mediators but also by their perceptual transparency. As stressed by Johansen, for Aristotle there is indeed a strict parallel between perceiving colors through a visual medium thanks to its transparency and hearing sounds through an auditory medium thanks to its resonance: “resonance serves the

4. As suggested by an anonymous referee of this journal, what seems to make a particular material apt to transmit sounds is the fact that it contains a fluid. The molecules of a fluid are indeed free to move and can therefore propagate sound waves. A similar idea is to be found in Heider (1959: 14–18).

5. For the view that according to Aristotle perceptual media not only include external entities causally involved in the perceptual process but also comprise the perceptual system itself, see Gregoric and Fink (2022: 31), Johansen (1997: 146, 187).

same function in the mediation of sound that transparency serves in the mediation of colour” (1997: 152).

For the purposes of this paper, I will focus on three main families of auditory media: materials, resonant objects—or resonators—and the spaces surrounding listeners.

Although these three families of auditory media are not exclusive and contribute to what we hear in very different ways, their role in audition, I shall argue, is basically the same and this role bears a fundamental relation to timbre. But before we can appreciate their shared features, let’s have a look at their specificities.

1) Sound perception is greatly determined by the materials through which sound waves propagate. To travel from their source to the listener’s ears, sounds waves must travel through a material medium like air or water. The propagation of sound waves 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 or sounding object), the sound wave is propagated through the medium, usually the air, by means of a chain of particle-to-particle interactions. Although we are always surrounded by auditory media, we seldom notice their presence, because like all perceptual media they are transparent and therefore imperceptible while functioning as media. In fact, we notice their presence only when their properties change or when we move from one auditory medium to another. Consider the case of a wall of bricks. Bricks absorb or reflect sounds with high frequencies but transmit those with low frequencies. This explains why a wall can protect you from your neighbors’ voices but not from the soundtracks of their superhero movies, which emit large amounts of low-frequency sound.

2) Acoustic materials are central to explanations of auditory perception because they fill the interspace between the source of the sounds and the listener, but also because objects that function as resonators are made of materials with different acoustic properties. Some sounds can be directly heard through air, but many sounds need to be amplified to reach our ears. Most musical instruments for example have a resonating body which can amplify the sounds produced by their vibrating strings or by a musician’s lips.

Yet, according to Aristotle’s view, objects that function as resonators must be considered as auditory media since they convey sounds to the perceiver. Unlike a sounding object which starts to vibrate at particular frequencies when stimulated by broad-spectrum energy—for example, by friction or by being struck—the vibrations of the resonator depend on the original vibrations. Like

air, a resonant object “does not make its own sound (it is not responsible for the sound) but *resounds* to the sound of something else. . . . The resonant is that *through* which the sound of another thing can be heard” (Johansen 1997: 152–53).

3) Another important family of auditory media is constituted by the space—enclosed or not—that surrounds the listener. A listener’s spatial surroundings comprise a great variety of elements, all of which have different acoustic properties and all of which contribute to the sounds we can hear. As expressed by Schafer’s (1977) useful notion of a “soundscape,” the auditory characteristics of a particular environment are caused by the fusion of a particular surrounding acoustics with the production of sounds in that particular space. Depending on their density, their materials, and their geometries, spatial environments give rise to particular aural architectures (Blessner & Slater 2007) that contribute to the sounds we can hear within them. These sounds can indeed be dry, clear, warm, full, muffled, crisp, or muddy depending on whether they are heard in deserts, forests, cathedrals, concert halls, or underwater.

Although materials, resonant objects, spatial layouts and architectures contribute in different ways to the listener’s auditory experience, they all “mediate the sound from the thing that has sound to the sense of hearing” (Johansen 1997: 149). In this sense, they all fall under the Aristotelian notion of an auditory medium which is what “is responsible for the sound being audible to the perceiver” (Johansen 1997: 151).

As we have seen, the Aristotelian notion of an auditory medium functions as a causal link between the sound produced by a distal object and the perceiver. But its role is also phenomenological since it allows the sound to be heard *through* it. The role of a medium is indeed to take on the sound of something else and to transmit it without interference. For a sound to be heard through the medium, it must therefore be auditorily transparent or “transsonant” (Johansen 1997: 153). Take, for example, a string telephone, a children’s toy made by connecting the bottoms of two paper cups with a tautly held string. This simple device makes it possible to have a conversation with someone over distances of up to 100 feet without shouting. Because the cups and the string have a higher density than air, sound vibrations do not tend to peter out before they travel far as they 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 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 person speaking into the cup. The cups and the string are in themselves perfectly inaudible.⁶

Yet, hearing through air, through water or by using a string phone are all phenomenologically different. So how can we account for these differences if, as I have claimed, the auditory medium is transparent?

The answer can be found in Aristotle's claim that the perception of a perceptual medium is only derivative. Here is a famous passage from *De Anima* (418b4–6): "I call transparent that which is visible not, strictly speaking, in itself but because of the colour of something else." The medium is visible only insofar as a colored body can be perceived through it. Water and air have no color on their own he maintains. They are only colored in a "derivative way".⁷ The same idea can be extended to auditory media: auditory media have no audible qualities on their own, they are audible only in a derivative way, that is, by hearing the sounds they transmit. As I will argue, timbral properties perfectly fit this model. A violin's body or a concert hall have auditory qualities only in a derivative sense. Unlike sounding objects, their characteristic timbres are not due to how they produce sounds but rather to how they transmit the sound of other objects to the perceiver.⁸

But before we can examine in more detail what timbres are and how they relate to auditory media, we need to understand the relations between the Aristotelian account of auditory media and other standard philosophical views about sound and audition.

6. They are inaudible, provided that they function as auditory media. If the string between the cups is plucked and starts to vibrate freely according to its own intrinsic properties, it would become a sounding object and stop being an auditory medium (see §5). Similarly, due to its specular properties and its capacity to reflect light, a pane of glass can be visible. In that case, the visible region of the pane of glass stop being a visual medium and becomes a visible object (see Mizrahi 2018: §2.2).

7. Here is the passage where Burnyeat explains very eloquently what it means for a visual medium to be colored in a derivative way:

The transparent does not have within itself the cause of its visibility. That is to say, it has no colour of its own; it is not coloured in the way the visible object is. When it is illuminated, it is visible and coloured in a derivative way, thanks to the presence of a colour which belongs to a body. But what is the meaning of 'coloured in a derivative way'? Here is a very simple reply. The transparent is coloured in a derivative way when the colour of a body appears through it.

When the medium is actually transparent (*diaphanēs*), i.e. when the medium is such that colours can appear through it (*phainesthai dia*), they do appear through it.

At the same time, the transparent itself, the light, becomes visible in a way and coloured in a way—without being really coloured and, in consequence, without undergoing a real alteration. This non-real alteration—a quasi-alteration I shall call it—of the transparent consists in the fact that colours appear through it. (Burnyeat 1995: 412)

8. The case of musical instruments is somewhat complex since they are usually composed of two main parts: a sound source (strings, reeds, . . .) and a resonating body. As we will see, it is its resonating body that gives the instrument its characteristic color.

4. Audible Objects versus Auditory Media

Following a standard view in science, some philosophers identify sounds with disturbances in a medium. Although they recognize that sounds originate from a vibrating object, they distinguish the source of a sound from the sound itself, which they usually identify with the waves that travel through the medium (Perkins 1983; Sorensen 2009) or with patterns instantiated by sound waves (Nudds 2009a). But one reason to distinguish between sounds and their sources is precisely the observation that the medium affects the qualitative properties of sounds we hear. Yet, identifying sounds with some property of the medium contradicts the Aristotelian view of perceptual media sketched above and leads to the abandonment of the very notion of a medium.

As we have seen in §3, Aristotle introduces the notion of a perceptual medium to explain how perception at a distance is possible. He recognizes that the remoteness of perceptual objects from the perceiver requires the presence of a causal intermediary, the external medium, connecting the remote object to the sense organ. Now, in order to serve as a mediator capable of transmitting information from the perceived object to the sense organ, the causal intermediary cannot interfere with the information transmitted. The medium must be transparent to allow other objects to be perceived *through* it. In fact, as Aristotle argues, for a substance or object to work as a medium, it cannot be perceived in a non-derivative way.

This point is fundamental when considering the nature of sounds and the role of sound waves for audition. 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 source, which corresponds to the vibrating object that causes the medium to vibrate. This *medial* approach⁹ to the nature of sounds therefore conflicts with the Aristotelian view of perceptual media, which considers the medium to be transparent and only indirectly perceptible. In fact, if sounds are sound waves, as claimed by the medial approach, the medium is not a mere mediator but a directly audible object and there is thus 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 in fact consider audition to be distal and can therefore dispense with the existence of an external medium.

I believe that the wave theory of sounds is right to stress that the medium plays a fundamental role in auditory perception but that it is wrong to identify

9. See Casati and Dokic (2010) for an overview of medial theories of sounds.

the auditory medium with the audible objects. As the wave theory recognizes, the auditory medium shapes audition. But unlike audible objects or events, auditory media don't have audible qualities. However, although directly inaudible, auditory media change our auditory perception by changing our access to sounds. Consider the difference between hearing a violin playing in an open space and hearing it in an enclosed space. Because of reverberation, the sound of a violin in an enclosed space, like a concert hall, can be perceived from much farther away than in an open space. Concert halls provide each member of the audience with an almost ideal vantage point for listening to the sounds emitted from the stage. This remarkable achievement would have been impossible without the particular elements and materials that are put together to form precise architectures dedicated to creating the best musical experience possible (Beranek 2004). A concert hall is indeed a complex auditory medium built to transmit particular sounds, but it is also designed to conceal others. Like all perceptual media, auditory media shape perception by selecting the particular sounds to which they give access. As a telescope gives access to distant visible objects, a concert hall gives access to distant sounds that are otherwise inaccessible.

Auditory media shape our experiences by modifying the distance at which we can detect sounds, but they also select the particular frequencies that can reach our ears. In general, textiles like rugs and drapes prevent high frequencies from reverberating. Transmitting mostly bass sounds, environments that make use of such textiles create warm atmospheres devoid of overly bright sounds. But selecting the range of frequencies is not only a matter of comfort. As biologists emphasize, adaptation to the surrounding auditory medium is also a matter of survival. Because sounds propagate differently in different natural environments, the acoustic environment has an important impact on the evolution of most animals. For example, many species rely on acoustic signals for communication, but their communication is effective only by virtue of the specific constraints imposed by their environment, and these constraints therefore determine the evolution of their hearing systems and vocal apparatus. Our hearing apparatus is in effect restricted to a frequency range that determines the sounds we can perceive. In fact, the hearing system is itself an auditory medium that filters the sounds to which we have access. As Massin argues (2010: 103), there is a continuum between the perceptual system and the perceptual media located outside the perceiver's body:

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?¹⁰

The hearing apparatus can be therefore considered an auditory medium selected to promote successful reproduction and survival in a complex acoustic environment (Morton 1975; Wilkins et al. 2013).

Although we can hear La Callas in an open space, in a concert hall, through the radio or with hearing aids, we don't hear the empty space, the concert hall, the radio or the hearing aids while listening to La Callas sing. The extended notion of auditory medium presented in §3 captures all these different cases and allows us to understand what they have in common and how they shape auditory perception while remaining themselves inaudible.¹¹ Auditory reality is complex, because almost every surface, material, or object located between the sound-generating event and the observer—including the observer's own auditory apparatus—has an impact on the sound that is transmitted to the listener. Although I will continue to use simple nouns, such as “concert hall” or “radio”, to refer to the causal intermediaries in audition, one should always bear in mind that the causal process which explain hearing is in fact extraordinarily complex, because the sounds we hear critically depend on the presence of a multitude of elements with different acoustic properties. The auditory medium is not audible, as the wave theory of sounds maintains, but it nevertheless shapes our perception of sounds. We don't hear the carpet that covers the tiles, the flat and reflective ceiling, the marble sculpture, or the heavy velvet curtains, but all these elements determine our auditory perception by selecting the sounds that reach our ears.

Now that we have a better understanding of auditory media and how they modify our experiences by selecting the sounds we can reach, it is time to clarify their relations to timbres and the roles they play in the perception of timbre.

5. The Nature of Timbre

Timbres are often considered to be qualities of sounds. Along with loudness and pitch, they are supposed to distinguish sounds from each other. According

10. The view that perceptual media include not only external entities causally involved in the perceptual process but also comprise the perceptual system itself was in fact anticipated by Aristotle in *De Sensu*, where he argues that the eye has to be composed of a transparent matter in order to be functionally similar to the external medium (see Gregoric & Fink 2022: 31; Johansen 1997: 187).

11. Extending the concept of “auditory media” in this way also stresses their common characteristics with perceptual media in other sense modalities. Like auditory media, visual media, such as air, glass and water, are indeed transparent. Moreover, it can also be argued, following common sense, that photography, television and films are transparent and can be considered as genuine visual media in Aristotle's and Heider's sense (see Mizrahi 2021).

to this view, sounds are the bearers of timbres. But as §2 suggests, timbres are better conceived of as properties of sound sources. O'Callaghan argues that the absence of invariants exposed by psychoacoustic experiments (Handel 1995: 441) supports the view that timbres are determined by the sounding objects rather than by particular features of sounds (O'Callaghan 2007: 89). Similarly, Kulvicki argues that perception of timbres reveals that sounds are stable dispositions of objects to vibrate rather than transient properties or events:

We distinguish the sound of a guitar from the individual note it happens to play. . . . These results reflect precisely the kind of auditory constancy one would expect if sounds are stable dispositions of objects to vibrate in response to being thwacked. They suggest that auditory perception concerns not how some object happens to be vibrating at any given moment but rather how that object is disposed to vibrate across modes of stimulation. (Kulvicki 2008: 11)

Like O'Callaghan, Kulvicki identifies the bearers of timbres with objects. But unlike O'Callaghan, he denies that this identification entails a contrast between audition of sounds and audition of sources. According to him, audition is in both cases directed towards objects—not towards events or sound waves.

I agree with Kulvicki and O'Callaghan that timbres often concern objects. We refer to timbres in the first place to distinguish objects and the characteristic sounds they cause. Violins, clarinets, and voices have different timbres. By contrast, we rarely describe events as having timbres. Explosions, buzzings, scratchings, and meows correspond to characteristic sounds or successions of sounds, but do they have characteristic timbres? Is this contrast important? Does it tell us something about the nature of timbre? I believe it does, but not for the reasons usually brought forward.

According to Kulvicki, timbres are properties of objects because they reveal the stable dispositions of objects to vibrate in a characteristic way. For example, our capacity to recognize voices despite the auditory variations necessary for talking seems to support the view that “what we hear is rather complex, but it includes the sound of the vocal cords as well as the sound of the rest of the person, even though the primary stimulus for the latter comes from the sound made by the former” (Kulvicki 2008: 9). For Kulvicki, timbres are attributed to objects because sounds themselves are properties of objects. Unlike most widespread contemporary views, which identify sounds with events or sound waves, Kulvicki defends the view that sounds are in fact properties of objects.

Kulvicki's view is attractive because it supposedly shows how timbres can characterize both sounds and their sources. Indeed if sounds are stable dispositions of objects, this provides a straightforward explanation why timbres, conceived as

qualities of sounds, can reliably be attributed to objects and not only to the sounds themselves. Although this view is plausible for some objects, it is unconvincing for most objects. For example, bells and tuning forks, which cause a characteristic sound when thwacked, seem to support the view that sounds are dispositions of objects to vibrate at particular frequencies in response to stimulation. But this approach does not seem to provide a plausible account of voices. As Kalderon rightly stresses, people produce characteristic sounds they don't in fact have:

When speaking, I produce the sounds that I do by an internal activity driving the vibration of special parts of myself. Are these not sounds that I make but do not have? If the sound of my voice is something that I make but do not have, then its being the constant element in an auditory experience provides no reason for thinking that sounds are stable dispositions to vibrate when thwacked since these are sounds that bodies were meant to have rather than make. (Kalderon 2017: 90)

This point holds also for musical instruments. Violins have characteristic timbres that we recognize by hearing the sounds they make, but they don't have characteristic sounds. Although a violin maker can appraise a violin by tapping the wood, this is not the sound it produces when played. Unlike the timbre of a bell, the timbre of a violin cannot be identified with its disposition to vibrate in a characteristic way when thwacked. Contrary to Kulvicki's suggestion, stable dispositions of objects to vibrate in characteristic ways cannot, by themselves, explain the nature of timbres.

As Handel points out, timbre recognition is a very complex phenomenon that depends on the perception of a "rich" set of exemplars (Handel & Erickson 2004: 591). Rather than extracting timbral properties from their experience with singular sounds, subjects identify sound sources and their timbre by becoming acquainted with perceptual invariants through hearing sounds at different pitches or loudness levels (Handel & Erickson 2004: 608). Contrary to Kulvicki's proposal, timbres are not experienced as belonging to objects because they are audible features of the sounds these objects make. They are rather attributed to particular objects because the sounds of these objects form a special set of sounds that are united by specific invariant relations. As Handel stresses, "The cues that determine timbre quality are interdependent because all are determined by the method of sound production and the physical construction of the instrument" (1995: 441).

Rather than thinking of timbre as adhering to objects through the sounds they produce, we should understand timbre as an attribute of objects that arises from the unity they impose on the sounds we hear.¹² Timbres are indeed prop-

12. The relation between timbres and the unity they impose upon a particular set of sounds is discussed in §6.4.

erties of objects, because all the sounds we hear as coming from these objects are determined by the physical structure of these objects and in particular their role as resonators. When the vibrations of one object cause another object to vibrate, the second object is said to resonate with the first object. Unlike a sounding object, such as a bell, which starts to vibrate at particular frequencies when stimulated, the vibrations of the resonator depend on the original vibrations. Although a resonator does not cause vibrations, its reaction to existing vibrations is determined by its own physical properties, such as its shape or material makeup.

Resonators are important for timbre because they act as filters in auditory perception. As Stumpf (1926) discovered, the notion of formants is central to understanding the constancy of voices and instruments across variations of pitch and loudness. Formants correspond to a range of frequencies where there are peaks, or local maxima, in the spectral envelope of the sounds at different pitches. Yet, formants result from the physical structure of the resonator. Violins with similar shape and construction features have a similar set of formants and therefore sound alike. Although a violin's sounds are produced by the friction of its strings, it is only through the process of its body's resonance that a violin produces its characteristic sounds. The mere vibration of its strings would be dull and almost inaudible.

Resonators and their formants are central to timbres, but unlike proper sounding objects, whose vibrations are conditioned internally, resonators don't vibrate freely according to their own intrinsic properties. Resonators are mediators in Heider's sense, because their movements are determined by the vibrations which impinge on them and are forced on them from outside. Heider explains:

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.

Free vibration is a unitary event. It is released by an external influence and then takes its course. The process which follows the striking of a tuning fork does not consist of parts that are independent of each other. With a forced vibration one wave could be missing, if the external influence were interrupted for a short time, and the vibration then start anew. With a free vibration, the dropping out of a part is impossible. (1959: 7)

Consider again the case of violins. A violin is a very complex object, but we can distinguish two main components: the strings, which produce vibrations, and the main body, which is a resonator whose main job is to amplify the original sounds produced by the strings. In effect, depending on the frequency, the

different parts of the body will resonate with the strings' vibrations and amplify these vibrations by moving a larger quantity of air than the strings alone. Now, in doing so, the body enhances the different frequencies produced by the strings by different amounts and imposes in this way its own formants on the sounds we hear. As with the timbres of voices and other instruments, the timbre of a violin is explained by the way it transmits certain frequencies more efficiently than others, giving the instrument its distinctive color.¹³

To sum up, timbres are properties which objects have in virtue of the way they transmit sounds. They are not qualities of sounds. Although most instruments combine sounding objects that vibrate freely and resonators whose vibrations are externally conditioned, what gives the instrument its distinctive timbre is its role as a mediator. As the next section argues, although perceptual media function as intermediaries in perception, their role is crucial and has a great impact on our access to reality.¹⁴

6. Auditory Media as Timbre Bearers

Although timbre is generally conceived as a quality of sounds, it is better viewed as a property of auditory media in the Aristotelian sense explained in §3. I believe this approach has many advantages. First, it accounts for the frequent attribution of timbres to objects that don't have a sound of their own. Second, timbres are attributed not only to ordinary objects, like instruments, but also to surrounding spaces and architectural structures. This fact is elegantly accounted for by a theory that views auditory media as the bearers of timbre. Finally, conceiving auditory media as the bearers of timbre provides an original solution to the timbre-constancy problem and offers a new perspective on the way we conceive relationships between sounds. Let's have a closer look at these points in turn.

13. See Schneider (2018: 706): "The size and shape of the resonator largely determines the spectral envelope . . . which is sensed as a peculiar sound color. . . . Further, the same principle can be applied to families of instruments that cover different registers (soprano, alto, tenor, baritone, bass) but share the same basic sound color".

14. One interesting consequence of the view of timbre defended here is that objects or instruments, like bells or tuning forks, which do not have a resonating body and do not produce a variety of sounds, do not have any timbre, properly speaking. Bells, triangles and many percussion instruments differ indeed from pitched instruments in that they do not vary in pitch. The sounds produce by a bell are therefore not different in nature from the sounds produced by any mundane object that makes a characteristic sound when thwacked. Although percussions and most everyday objects don't have a characteristic timbre, they produce characteristic sounds that can be used by percussionists and composers to extend the sonic texture of their music. Thanks to an anonymous referee of this paper for raising this point.

6.1. *Objects*

As stressed in §2, timbres are commonly attributed to objects. We refer to the timbre of a violin or the timbre of a singer's voice and commonly individuate timbres by referring to the objects with which they are usually associated. Yet, if timbres are properties of objects, they cannot be directly accounted for by philosophical theories that identify sounds with properties of waves or events. According to these theories, timbres—as properties of objects—are properties of sources but not of sounds. By contrast, if sounds are properties of objects, as advocated by Pasnau (1999) and Kulvicki (2008; 2014), attributing timbres to objects seems quite straightforward, because timbres as sound qualities would inhere in objects. The problem, however, is that the notion of timbre is intrinsically tied to sound variations and not, as this view suggests, to sound stability. We recognize the timbre of an instrument by hearing the different sounds it produces, not by hearing a steady sound with stable qualities. The sounds made by a violin can change in pitch and volume while its timbre remains the same.

Moreover, the sounds of objects vary according to the mode of stimulation. A violin doesn't produce the same sounds if its strings are bowed or its body knocked, making it impossible to individuate its characteristic sound. As Isaac explains, the musical use of violins for individuating its proper sounds seems quite arbitrary:

I think examples such as this reveal just how heavily the stable dispositions view relies on musical examples for its plausibility; when we speak casually of the sound or timbre of a violin, the kind of consideration that motivates Kulvicki, it is really shorthand for the sound or timbre of a violin played in the usual manner—without reference to our usual mode of interaction with the object, a reference implicit in musical instrument examples, there is no way to identify the timbre of the object. (Isaac 2018: 516)

Although Kulvicki's approach seems unable to explain why instruments have characteristic timbres, it rightly emphasizes the importance of the physical constitution of objects for their distinctive timbre. We may not have cracked Stradivari's secrets, but we know that the exceptional timbres of his violins reside in their physical construction. The problem therefore seems to be how to reconcile the fact that objects don't have proper sounds with the fact that their physical makeup explains their timbre.

According to the auditory-media approach proposed here, material bodies can act as mediators in audition without being the bearers of sounds or even participating in sound events. Like vision, audition extends to distant objects and relies on intermediaries to carry information from these perceptual objects to the

perceiver. Yet, to transmit information about perceptual objects, a medium must be transparent and not interfere with the information transmitted. This is the reason auditory media don't have audible qualities in the strict sense. Auditory media transmit sounds without being heard, but they do so differently according to their particular physical properties. What is required to transmit a sound is that the initial vibrations corresponding to the sound we hear are propagated through a medium until they reach the listener. Now, depending on the materials, sound-wave propagation will vary in speed, frequency, and direction and therefore shape in its own way how we access the world. A wall that does not transmit high frequencies will transmit only low-frequency sounds; a medium like water, which is denser than air, will transmit sounds faster and farther but will make it difficult or even impossible for human subjects to localize them. All auditory media transmit sounds in a particular way and therefore impose their characteristics on auditory experiences. Although they are not audible, auditory media determine a causal perspective on the auditory information available to the observer. As with spatial perspectives, which delimit what is perceptible and not perceptible from the observer's standpoint, auditory media affect the phenomenology of auditory experiences because each particular medium offers a limited and different form of access to the auditory world.

The answer to my initial question, "How does the physical makeup of objects contribute to their timbre if they don't have sounds?" runs, then, as follows. As we have seen, most instruments and many ordinary objects are not only sources of sounds but also mediators that select the sounds we can hear.¹⁵ Yet, each mediator's selection of sounds is singular, because it imposes its own physical characteristics on the sounds we perceive through it. Although we don't hear the mediator, its timbre is nonetheless accessible through the unique set of sounds it selects.¹⁶

6.2. *Surrounding Spaces and Architectural Structures*

The fact that any auditory experience takes place through an auditory medium (or a particular combination of media) that delimits what is perceptually accessible to the observer becomes manifest when the medium varies. The cases of "empty" spaces and architectural structures provide particularly good examples. As Blesser and Salter remark, a good way to demonstrate that we are always aware of the environment in which auditory experiences take place is to displace familiar sounds to unfamiliar environments:

15. The distinction between sound sources and sound mediators has important similarities with the source-filter model proposed by Handel (2006) for timbre perception.

16. For more details see below and §6.4.

Transported to an open desert, urban traffic would not have the aural personality of a dense city environment. Moved to a forest, a symphony concert would not have the aural impact, intimacy, and immediacy of a concert hall. Nor could the aural personality of singing in the bathroom, which takes advantage of the resonances of small spaces, be duplicated in a large living room. In each contrasting space, even if the sound sources were to remain unchanged, the aural architecture would change. (Blessner & Salter 2007: 2)

Philosophers and psychologists recognize and discuss awareness of space through audition, but few acknowledge that spaces don't have intrinsic audible properties like sounding objects or events. Although we can hear auditory variations caused by the geometry of the surrounding space and its architectural components, these elements don't have intrinsic audible features. They can be experienced only by hearing the sounds of other objects or events.¹⁷ To dissolve this apparent paradox, let's return to the case of musicians performing in a concert hall. Although we don't hear the concert hall in the same way we hear the musicians playing on the stage, there is a clear sense in which the architectural specificities of the concert hall are accessible through the musical performance we hear. Consider Karajan's testimony about his experience when playing at the Stadt-Casino in Basel: "This is a typical rectangular hall—small, with a wonderfully clear and crisp resonance. It is almost perfect for Mozart" (Beranek 2004: 461). Because sounds depend on a medium to be transmitted from their distant location to the listener, the specificities of the media are accessible through the way they transmit sounds. Concert halls are particularly relevant to our discussion of spaces as auditory media because a concert hall has a very specific auditory purpose, which is to serve as an "extension of the musical instruments played within it" (Blessner & Salter 2007: 135). Enclosed spaces like concert halls store the sounds they contain and distribute them throughout the space before they gradually fade away. Like the body of a violin, their purpose is indeed to capture and make accessible to listeners all the minute variations contained in complex musical works. As Blessner and Salter (2007) stress, resonant enclosures like concert halls are integral components of the musical performance. Musicians must familiarize themselves with and learn to play in response to the specific acoustics of concert halls as they learn to play their instruments. Conductors in particular must adapt and master the special acoustics of a particular venue. As Maestro James De Priest explains, this is one of the reasons each orchestra develops its own unique color:

17. For a defense of the claim that spaces are audible, see Young (2017).

A few years ago I had the chance to compare the effects of two very different acoustics in two halls on the playing techniques of their home orchestras—Boston Symphony Hall and Philadelphia Academy of Music. The Boston Symphony Orchestra, playing in its legendary Hall, offered a Brahms 2nd Symphony that was lean, angular and sharply etched with splendid clarity: a crisp and highly articulated Brahms, which was ideal for the Hall's long reverberation time. A great hall tends to be permissive and inviting in the process of music making. A few months later I conducted the Philadelphia Orchestra in the Academy of Music. As with Boston, I invited the orchestra to repose into the Brahms of their tradition—a more rounded and expansive performance. The results were shaped as much by the acoustics of the hall as by any other factor. Both of these magnificent orchestras are products of their environments as their respective sonic profiles witness.

As with every orchestra, the Philadelphia developed its sound in response to the acoustics of its home, and although some might view the blanketing effect of the Academy as a drawback, it has in fact demanded of its musicians, lushness, a distinctive expansiveness, that would never have been realized in a more forgiving hall. (Beranek 2004: 6–7)

Although we can't hear spaces and their particular layout, they shape what we hear by virtue of their unique way of selecting the sounds we hear through them. Without the reverberation of an enclosed space, all the subtleties of classical musical compositions would be lost for distant listeners. The sounds heard without reverberation would be weak, dull, muffled, and almost imperceptible. Reverberation allows sounds to travel with ease. Isaac Stern puts the point nicely, "Reverberation is of great help to a violinist. . . . The resulting effect is very flattering. It is like walking with jet-assisted takeoff" (Beranek 2004: 7).

Although they are inaudible, spaces have their own acoustic qualities. Whereas cathedrals amplify and blend successive sounds, creating an enveloping reverberation and destroying most aural localization cues, smaller places, like hotel bars, offer warm and intimate atmospheres by restricting reverberation and keeping conversations private by preventing sound from passing from space to space. Spaces do not need to be heard to have intrinsic acoustic qualities, as Young (2017) argues. And their acoustic specificities do not need to be reduced to the timbral qualities of the sounds we hear. As I shall explain below, timbres are not properties of sounds, and they therefore do not vary as sounds do.

6.3. *Timbre Constancy*

We perceive the timbre of a particular instrument¹⁸ as constant despite changes in the sounds it produces. This phenomenon of timbre constancy constitutes a challenge for most theories of sounds because it is unclear what remains constant when sounds vary. What distinguishes a Stradivarius from an ordinary violin is not how it produces a particular sound or how great its projective power is, but rather how incredibly varied and subtle the shades of its tones are and how responsive it is in the hands of a skilled musician. Although we identify timbre by hearing sounds, explaining timbre constancy in terms of some specific feature of the sounds we hear is problematic, because as Isaac (2018) explains, individuating timbre relies not on one phenomenon but rather many phenomena that are different in nature.

To identify a particular timbre we need to do more than hear a particular sound, we need for example to recognize how different sounds follow a particular dynamic pattern of attack and decay. As rightly pointed out by Isaac (2018: 515): “one problem is that timbre identity is constitutively tied to changes in a sound over time.” This is true for instruments but also for the acoustic environment where those objects are heard. Although we can recognize and appreciate La Callas’s voice in different places, variations in the acoustic properties of the environment are accompanied by changes in what we hear. The empty space that surrounds a listener seems indeed to have its own timbral properties as attested by Blesser and Salter’s summary of Beranek’s characterization of concert halls:

Using a lifetime of professional experience, Beranek sorted concert hall quality into 18 distinct concepts, each of which was represented by a word or phrase that served as a label for a perceptual experience. Some of the more prominent labels included “intimacy,” “fullness of tone,” “clarity,” “warmth,” “brilliance,” “balance,” “blend,” “ensemble,” and “texture.” (Blesser & Slater 2007: 218)

Like the timbre of instruments, the timbral characteristics of the environment are not characterized by some constant features of the sounds heard, but rather by some invariance among the diversity of sounds they transmit. As Blesser and Slater explain, we don’t directly hear empty spaces. Our awareness of acoustic spaces is dependent on our perception of the sounds they contain: “Just as light sources are required to illuminate visual architecture, so sound sources (sonic events) are required to ‘illuminate’ aural architecture in order to make it aurally

18. As we have seen throughout the paper, musical instruments are complex objects which usually comprise two main parts: a sounding source (strings, reeds, . . .) and a resonating body. As we will see, it is its resonating body that gives the instrument its characteristic color.

perceptible” (Blesser & Slater 2007: 15–16). For example, the intimacy of a room or a concert hall is due to the fact that the sounds appear close to the listeners. The space itself need not be small; its intimacy arises from the shortness of the delay between the direct sound and the onset of its first reflection (Blesser & Slater: 218). The clarity of a concert hall corresponds to its capacity to keep sounds separated, whereas its fullness corresponds to its capacity to extend the attack and release of each tone.

Different timbres involve therefore different perceptual invariants, rather than different fixed qualitative characters of the sounds we hear. What remains the same when listening to the sounds produced by a clarinet or a cello is indeed a structure that remains invariant when notes that vary in pitch are played with each particular instrument. Moreover, what is remarkable is that we recognize the timbre of each particular instrument despite the fact that “each instrument changes acoustically and perceptually in different ways” (Handel & Erickson 2004: 589). Similarly, what remains the same when listening to sounds in a park or in a concert hall is not some particular quality of the individual sounds we hear, but rather the fact that different sounds played with different instruments in a particular environment exhibit a perceptual constancy. By dissociating audible objects and events from the bearers of timbre, the view of auditory media presented here provides an original approach to timbre constancy and explains why timbre perception is associated with a constancy exhibited by variations of sounds rather than with some qualitative feature of the sounds themselves. It also explains why timbre experiences are multilayered and complex. The ability to identify timbre depends indeed “on a ‘rich’ set of exemplars” (Handel & Erickson 2004: 591). Since timbres correspond to invariants across changes of sounds, to identify a particular timbre, the listener must be presented with a relevant set of exemplars. And since most auditory experiences take place through multiple auditory media, such as an instrument (its resonating body), a filling material and a particular surrounding space, attributing timbral properties to these different components can only be done if the set of sounds available in these experiences display the kind of variety necessary for the subject to discriminate between them. For example, although air and the particular geometry of a concert hall both contribute in a particular way to the sounds we can hear, their respective contribution—and therefore their respective timbre—may not be individually accessible to the listener if the kind of sound variations that disentangle them do not occur.

6.4. *Sound Families*

According to the approach defended above, timbre is not a quality of particular sounds but a property of auditory media. Each medium selects the set of sounds

that travel to our ears and thereby imposes its own characteristics on auditory experiences. According to this view, what characterizes timbre is therefore not a single feature or property of sound sources or sounds themselves, but rather a set of sounds brought together by particular relations of resemblance and exclusion.

What brings together all the sounds of a particular Stradivarius or all the sounds heard in a cathedral is the fact that they constitute different “families of sounds.”¹⁹ What explains the phenomenological variations due to the environment is the special perspective associated with a particular medium. Reverberation in a cathedral gives access to distant sounds but diminishes our capacity to locate them with precision. By contrast, rooms filled with upholstered furniture reduce the distance and the amount of time necessary for sounds to become inaudible, but they better preserve the spatial information they carry. Although the sounds produced remain the same, our perception of them can vary because our perceptual access to them can vary according to the auditory medium. Timbre constancy does not, therefore, correspond to our capacity to track audible qualities of sounds or qualitative properties of their sources; it corresponds to the unity imposed on our auditory experiences by a particular auditory medium. Understanding timbre as a feature of auditory media offers a new perspective on the way we conceive relationships between sounds.

A Stradivarius is considered superior because it offers a palette of sounds different from the palette of sounds produced by a cheap violin. Good violins are powerful and radiant—able to elevate their sounds above whole orchestras and reach the listeners without fading. Moreover, their responsiveness and balance make them apt to respond promptly and with precision to the complex and subtle variations of the violinist’s movements. A violin that can transmit all the variations of any piece of music must be extraordinarily versatile and therefore produce as large a palette of sounds as possible. Although there are not poor and bad auditory media per se, some media are better suited to particular goals than others. Sometimes it is good to transmit a large palette of sounds, but sometimes it is not. For example, extracting information in a noisy environment can be improved by filtering out sounds of particular frequencies. In that case, choosing an auditory medium characterized by a smaller palette of sounds is judicious, because it can eliminate perturbations caused by unwanted sounds. Focus and expressivity often require different media. Timbres are characterized by the palette of sounds they can transmit. The sounds of these palettes vary, as do their sizes. The size of the set of sounds that characterizes an auditory medium matters because an extended set allows both more nuances and more contrasts. By contrast, auditory media that transmit only a restricted set of sounds allow

19. For the idea that visual systems and viewing circumstances determine unique color families, see Kalderon (2007) and Mizrahi (in press).

fewer nuances and contrasts, consequently restricting the listener's discriminating capabilities.

Interestingly, color perception exhibits a similar contrast. The set of colors we can perceive at a time changes according to the nature of the light. The colors we can perceive in daylight—a full-spectrum light—are more vibrant and varied than those we can perceive when the light is restricted to a narrow wavelength band. The colors we perceive under an illuminant composed of long wavelengths only are all reddish. We don't see any blue or yellow surfaces. By contrast, surfaces perceived under a short-wavelength illuminant are bluish, but not vivid red or yellow. Although it is useful to restrict the wavelengths of an illuminant in certain endeavors—such as discovering hidden materials or molecules—we tend to prefer full-spectrum illuminants because they give us access to more color nuances and contrasts. Our preference for daylight is justified not because it reveals an object's real color, as Allen (2010) argues, nor is it merely arbitrary, as Cohen (2008) argues. Natural daylight is generally preferred because it provides a rich palette of colors that allows us to easily discriminate between surfaces and identify objects.

I have argued that auditory media select the audible features accessible to listeners. The sounds perceived through a particular medium therefore share particular features. They constitute distinct families “that display some unity despite their plurality” (Kalderon 2021: 328). All the sounds coming from a clarinet are related by a set of similarities and differences they do not share with sounds belonging to other families, like the family constituted by the sounds coming from a violin. The set of resemblance relations that holds between the members of one family doesn't hold between the members of other families. As we have seen in §6.3, another way to understand the unity exhibited by timbres is through the notion of an invariant (Gibson 1965: 68). Timbres are invariants in the sense that they correspond to what remains constant across a special kind of variation. Unlike pitch, timbres are not properties of particular sounds, rather they correspond to some constancy among a class of sound variations. The unity exhibited by timbres is therefore always relative to a subset of sounds—a particular family.²⁰

Timbres, like illuminations, are fundamental for perception because they unite sounds into families, which provide the necessary framework for gaining knowledge about reality. Auditory experiences vary according to the sounds we hear, but it is because the sounds transmitted by a particular medium are united

20. Since most of the time sounds are heard through a combination of auditory media, such as an instrument, air and a particular type of surrounding, they belong to multiple subsets of sounds and carry therefore quite diverse types of information. For example, when one hears la Callas, one perceives not only her voice, but also the special acoustic properties of the concert hall where the performance takes place.

by a set of resemblances and differences that those sounds exhibit properties independent of the sounds themselves. For example, it is the fact that at a certain location we hear sounds transmitted by a unique medium—or a unique combination of media—that allows us to evaluate their distance. If we were to hear sounds through continually changing media, it would be impossible to compare them spatially.

Although timbre provides information about reality, when it comes to music, the role of timbre is more complex and is not limited to carrying environmental information. As I will show next, since timbre is central to music, it is not surprising that the study of timbre and the study of music have always been closely connected.

7. Conclusion: Timbre and Music

The notion of timbre has always been associated with music and musical instruments as attested by this extract from Rousseau's Dictionary of Music:

In regard to the difference which is found also between sounds with respect to the quality of timbre, it is evident that it is due neither to the degree of elevation nor even to that of force. It will be in vain for an oboe to place itself in unison with a flute, it will be in vain to sweeten the sound to the same degree, the sound of the flute will always have a *je ne sais quoi* of mellow and sweet; that of the oboe *je ne sais quoi* of rude and sharp, which will prevent the ear from confounding them.²¹ (Rousseau 1780–89: 629)

Instruments and voices have been important to the study of timbre, because in contrast to most other objects, instruments have been created to exhibit very large palettes of sounds. Although instruments and voices have been important for the study of timbre, what is more significant is the role of timbre in music. As rightly noticed by Rousseau, what characterizes an instrument is not only the variety of sounds it produces, but also the fact that there is a qualitative feature — its timbre — that unites these different sounds and differentiates them from the sounds of other instruments.

Understanding timbre is essential for understanding western music, because it provides the unity from which variations of pitch can emerge. Schaeffer

21. Helmholtz has a very similar definition of timbre: "By the quality of a tone [timbre, Klangfarbe] we mean that peculiarity which distinguishes the musical tone of a violin from that of a flute or that of a clarinet or that of the human voice, when all these instruments produce the same note at the same pitch" (Helmholtz 1877: 10).

explains: “The function of the piano or the violin is to produce objects having enough characteristics in common (timbre) for their value (pitch) to be differentiated” (Schaeffer 2017: 238). Although most composers and writers on music have insisted on the dominant role of pitch in western music (Hindemith 1949; Toch 1977), there would be no melodies or musical structures if we were unable to group together the successive sounds having identical timbres. To understand the structure of a composition, it is indeed essential to understand the relationship between two or more melody lines that are played at the same time. This would not be possible if we could not group the successive sounds forming these melodies according to their timbre.

I have argued that timbre is a feature of auditory media. This view claims that, contrary to a widely held view, timbre is not an audible quality of sounds as are pitch and loudness. Timbre characterizes objects, materials, and spaces that propagate sounds but not the sounds themselves. If timbre, as I have suggested here, characterizes instruments as well as particular spaces, it is not surprising that some musicians and composers have explicitly integrated the performance space in their work. Rather than attempting to “de-emphasise the acoustic characteristics of particular auditoriums in order to attain a kind of acoustic norm appropriate to the kind of music being played” (Marshall 1976: 284), composers can choose indeed to assign an aesthetic role to the space where the music is played. Composers can choose for example to exploit the special acoustic of a space, like a church, to merge sounds in a particular way, but they can also relegate to the background the particular sounds that are heard in order to bring forward the special acoustic properties of the space where those sounds are played, as in Lucier’s famous *I Am Sitting in a Room* (1969).

Concert halls and churches can impact how music is perceived. But nothing has been more influential than the recent developments in miking, recording, electrical amplification and mixing for the way we listen to music. These new technologies not only reproduce existing spaces, but they create new auditory media *à la carte* in order to accommodate any music and any style (Doyle 2005). This evolution is particularly explicit in popular music and electroacoustic music where the music is inseparable from its production. As summarized by Brian Eno in an interview for *Artforum*:

Musicologists who say that everything pop musicians are doing was really known by about 1820 may be correct in terms of compositions written down on paper, but they ignore where the true innovation is taking place. The interest today isn’t in developing serial music or polyphony or anything like that. It is in constantly dealing with new textures. One of the interesting things about pop music is that you can quite often identify a record from a fifth of a second of it. You hear the briefest snatch of

sound and know, “Oh, that’s ‘Good Vibrations,’” or whatever. A fact of almost any successful pop record is that its sound is more of a characteristic than its melody or its chord structure or anything else. The sound is the thing that you recognize. (Korner 1986: 76)

Acknowledgments

Thanks to Kevin Mulligan and two anonymous referees for this journal for extensive and helpful comments.

References

- Allen, Keith (2010). In Defence of Natural Daylight. *Pacific Philosophical Quarterly*, 91, 1–18.
- ANSI (1960). *Psychoacoustic Terminology: Timbre*. American National Standards Institute.
- Beranek, L. (2004). *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*. Springer
- Blesser B. and L.-R. Salter (2007). *Spaces Speak, Are You Listening?: Experiencing Aural Architecture*. MIT Press.
- Bregman, Albert (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press.
- Burnyeat, M. (1995). Additional Essay : How Much Happens when Aristotle Sees Red and Hears Middle C? Remarks on *De Anima* 2. 7–8”. In M. C. Nussbaum, and A. Oksenberg Rorty (Eds), *Essays on Aristotle’s De Anima* (408–21). Oxford University Press.
- Casati, R. and J. Dokic (1994). *La philosophie du son*. Chambon.
- Casati, R. and J. Dokic (2010). Sounds. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2010/entries/sounds/>
- Cohen, Jonathan (2008). Colour Constancy as Counterfactual. *Australasian Journal of Philosophy*, 86(1), 61–92.
- Dickens, C. (1838). *Oliver Twist*. Gutenberg Project. <https://www.gutenberg.org/files/730/730-h/730-h.htm>
- Donnadieu, S. (2007). Mental Representation of the Timbre of Complex Sounds. In J. W. Beauchamp (Ed.), *Analysis, Synthesis, and Perception of Musical Sounds* (272–319). Springer.
- Doyle, P. (2005). *Echo and Reverb: Fabricating Space in Popular Music*. Wesleyan University Press.
- Gibson, J. J. (1965). Constancy and Invariance in Perception. In G. Kepes (Ed.), *The Nature and Art Motion*. Brazillier.
- Gregoric, P. and J. L. Fink (2022). Sense Perception in Aristotle and the Aristotelian Tradition. In J. Toivanen (Ed.) *Forms of Representation in the Aristotelian Tradition* (Vol. 1, 34–39). Brill.

- Handel, S. (1995). Timbre Perception and Auditory Object Identification. In B. C. J. Moore (Ed.), *Hearing* (425–61). Academic Press.
- Handel, S. (2006). *Perceptual Coherence: Hearing and Seeing*. Oxford University Press.
- Handel, S. and M. L. Erickson (2004). Sound Source Identification: The Possible Role of Timbre Transformations. *Music Perception: An Interdisciplinary Journal*, 21(4), 587–610.
- Heider, Fritz (1959). Thing and Medium. *Psychological Issues*, 1, 1–34.
- Hindemith, P. (1949). *Elementary Training for Musicians* (2nd ed.). Associated Music Publishers.
- Isaac, A. (2018). Prospects for Timbre Physicalism. *Philosophical Studies*, 175(2), 503–29.
- Johansen, T. (1997). *Aristotle on the Sense-Organs*. Cambridge Classical Studies. Cambridge University Press. <https://doi.org/10.1017/CBO9780511518461>
- Kalderon, Mark E. (2007). Color Pluralism. *Philosophical Review*, 116, 563–601.
- Kalderon, Mark E. (2017). *Sympathy in Perception*. Cambridge University Press.
- Kalderon, Mark E. (2021). Monism and Pluralism. In D. H. Brown and F. Macpherson (Eds.). *The Routledge Handbook of Philosophy of Colour*. Routledge.
- Korner, A. (1986). Aurora Musicalis. *Artforum*, 24(10).
- Kulvicki, J. (2008). The Nature of Noise. *Imprint*, 8(11), 1–16.
- Kulvicki, J. (2014). Sound Stimulants. In Ed. D. Stokes, M. Matthen, and S. Biggs (Eds.), *Perception and Its Modalities* (205–221). Oxford. University Press.
- Leddington, J. (2014). What We Hear. In R. Brown (Ed.), *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience* (321–34). Vol. 6 of Studies in Brain and Mind. Springer.
- Leddington, J. (2019). Sounds Fully Simplified. *Analysis*, 79(4), 621–29.
- Marshall, S. (1976). Alvin Lucier's Music of Signs in Space. *Studio International*, 192, 285–89.
- Massin, O. (2010). *L'objectivité du toucher* (doctoral dissertation). Aix-Marseille University. http://www.academia.edu/3010474/Lobjectivite_du_toucher
- Mizrahi, V. (2018). Perceptual Media, Glass and Mirrors. In T. Crowther and C. Mac Cumhail (Eds.), *Perceptual Ephemera* (238–59). Oxford University Press.
- Mizrahi, V. (2020). Recorded Sounds and Auditory Media. *Philosophia*, 48, 1551–67.
- Mizrahi, V. (2021). Seeing Through Photographs: Photography as a Transparent Visual Medium. *The Journal of Aesthetics and Art Criticism*, 79, 52–63.
- Mizrahi, V. (in press). Color Constancy Illuminated. *Dialectica*.
- Morton, E. S. (1975). Ecological Sources of Selection on Avian Sounds. *The American Naturalist*, 109(965), 17–34.
- Nudds, M. (2009a). Sounds and Space. In M. Nudds and C. O'Callaghan (Eds.), *Sounds and Perception: New Philosophical Essay* (69–96). Oxford University Press.
- Nudds, M. (2009b). What Are Auditory Objects? *Review of Philosophy and Psychology*, 1(1), 105–22.
- O'Callaghan, Casey (2007). *Sounds: A Philosophical Theory*, Oxford: Clarendon Press.
- Pasnau, R. (1999). What Is Sound? *Philosophical Quarterly*, 49, 309–24.
- Perkins, Moreland (1983). *Sensing the World*. Hackett.
- Rousseau, J.-J. (1780–89). Dictionnaire de musique. In *Collection complète des œuvres* (Vol. 9). Genève. <https://www.rousseauonline.ch/>
- Sankiewicz, M. and G. Budzynski (2007). Reflections on Sound Timbre Definitions. *Archives of Acoustics*, 32, 591–602.

- Schaeffer, P. (2017). *Treatise on Musical Objects: An Essay across Disciplines*. University of California Press.
- Schafer, R. (1977). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Knopf.
- Schneider, Albrecht (2018). Perception of Timbre and Sound Color. In Rolf Bader (Ed.), *Springer Handbook of Systematic Musicology* (687–725). Springer.
- Sorensen, R. (2009). Hearing Silence: The Perception and Introspection of Absences. In M. Nudds and C. O’Callaghan (Eds.), *Sounds and Perception: New Philosophical Essay* (126–45). Oxford University Press.
- Stumpf, C. (1926). *Die Sprachlaute: Experimentell-phonetische Untersuchungen nebst einem Anhang über Instrumental-Klänge*. Springer.
- Toch, Ernst (1977). *The Shaping Forces in Music: An Inquiry into the Nature of Harmony, Melody, Counterpoint, Form*. Dover.
- von Helmholtz, Hermann (1877). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, 4th ed. Braunschweig: Vieweg und Sohn. English tr. 1954 *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (2nd ed.). Dover.
- Wilkins, M. R., N. Seddon, and R. J. Safran (2013). Evolutionary Divergence in Acoustic Signals: Causes and Consequences. *Trends in Ecology & Evolution*, 28(3), 156–66.
- Young, Nick (2017). Hearing Spaces. *Australasian Journal of Philosophy*, 95(2), 242–55.