A brief argument for, and summary of, the concept of Sonic Virtuality

Grimshaw, Mark

Published in:
Dansk Musikforskning Online

Publication date:
2015

Citation for published version (APA):
Grimshaw, M. (2015). A brief argument for, and summary of, the concept of Sonic Virtuality. *Dansk Musikforskning Online*, *2015*(Special Edition), 81-98.
http://www.danishmusicologyonline.dk/arkiv/arkiv_dmo/dmo_saernummer_2015/dmo_saernummer_2015_lyd_musikproduktion_05.pdf

MARK GRIMSHAW

A brief argument for, and summary of, the concept of Sonic Virtuality[1]

Introduction

When that august body, the Acoustical Society of America, publishes a definition of sound in the American National Standards bible on *Acoustical Terminology*, and an innocent question on the wording of that definition to an acoustics-based colleague of mine provokes a negative response – "[the] definition [is] out of scope for most of the purposes I know [the definition] is only operational for some purposes [...] It is therefore necessary to use domain-specific definitions"[2] – then there is clearly something quite wrong. In extremis, one might reasonably ask, if there is no agreement as to what sound is even among acousticians, is the entire scientific basis for the study of acoustics fundamentally flawed?[3]

As this essay makes clear in its opening sections, there historically have been and currently are several definitions of sound and so among the first aims of the essay are to enumerate some of the most important definitions, to tease out differences between them, and to highlight inconsistencies both with each definition and in the context of our experiences, physically and phenomenologically, of what sound is.

What is the purpose of such an initially hermeneutic approach? It is precisely this: to argue for a definition of sound (and its underlying concepts) that is more consistent with recent research and thinking, that is more consistent with our phenomenological experience, and that has more use-value in the context of new technological developments. On this last point, my interest has been sparked by research collaboration with colleagues in the field of computer game sound. In particular, our work with biofeedback has, for our thinking about sound, its effects, and how to design affective sound, inspired and clearly necessitated the need for a new definition of what

1 In a 2015 book, my co-author Tom Garner and I put forward an alternate definition and concept of sound (M. Grimshaw and T. Garner, *Sonic Virtuality: Sound as Emergent Perception*, (New York: Oxford University Press, 2015)) that was in opposition not only to the standard view of sound (the one used in acoustics) but that also opposed other more philosophical definitions that have appeared over the last decade or so. I presented a brief presentation of the book and its ideas to the *Musik og lyd konference* in early 2015 and what follows is an expansion of that paper and so a summary of some of the major points of *Sonic Virtuality*.

2 Personal email communication, 1st September 2015.

3 Probably not. But, as I explain in this essay, acousticians *are* misguided concerning *what* they study. Thinking that what they study is sound, they are in fact studying sound waves.

sound is.[4] What becomes clear from this research (and other research investigating responses to audio) is the highly individualistic sound we experience – the phenomenon of sound is context-specific where that context includes the perceiver's brain and the wider environment. Rather than attempt to use statistical methods and rationales to smooth over such differences in order to produce a generalized, mass-produced, one-size-fits-all audio artefact (for computer games and elsewhere), I prefer the approach of sound design that celebrates and utilizes these personal differences because this, I believe, increases our engagement with computer games (and other multimodal artefacts), increases the possibility of enhanced presence in virtual environments, and leads to new sound design archetypes. Most importantly, though, such an approach, and the results deriving from it, gets us closer to what sound is as the object of design and as the object of study; in order to design sound, one must first know what it is. Hence the definition offered here.

I begin the essay proper by enumerating various current definitions of sound and discussing a number of identified problems with them. The first section deals with what is termed the standard definition of sound in the western world (there are actually several of these 'standard' definitions) and the inconsistent and incoherent uses to which this definition is put. Following this, I present some other definitions that are philosophical and, towards the end of the section, are more phenomenological. In this second section, I also present a short discussion of everyday listening. I then move onto a presentation of the new definition and a brief outline of the concept of *sonic virtuality* before closing with some speculative thoughts on the definition's and concept's use and implementation by sound designers.

The standard view of sound

The multiplicity of physics-based definitions

Whenever I ask my students *what is sound?*, beyond regular and frequent answers such as *noise*, *music*, or *speech*, almost invariably I am told that *sound is a sound wave*. This definition (or a variation of it) is probably the answer many of the readers of this article would provide because it is the standard Western definition as found in physics and specifically acoustics. The 10th edition of *The Concise Oxford English Dictionary* provides the following primary definition of the noun: "vibrations which travel through the air or another medium and are sensed by the ear." A fuller definition, in which I attempt to be more precise by dint of the use of jargon, may be stated as: *sound is a pressure wave moving through a medium and whose properties of direction of travel, frequency, and amplitude lie within the sensory range of an auditory system*.

The American National Standards Institute (ANSI) provides two (possibly three) definitions of sound to be used depending upon the context in which sound is to be

4 Biofeedback in this context involves the use and interpretation of data from psychophysiological devices such as electroencephalographic headsets for the real-time processing or synthesis of audio for the purposes of close control of arousal and emotions during computer gameplay. As an example: in a survival horror game, the game engine will sense that you are not scared enough and will immediately respond with more scary sounds.

used or studied. Sound is either "(a) Oscillation in pressure, stress, particle displacement, particle velocity, etc., propagated in a medium with internal forces (e.g., elastic or viscous), or the superposition of such propagated oscillation" or the "(b) Auditory sensation evoked by the oscillation described in (a)."[5] Sound is thus, and variously, an oscillation in any of an unstated number of physical properties, the superposition of such oscillations (and one must assume the plural here), or the auditory sensation of this oscillation (presumably also the sensation of superpositions of such oscillations).[6]

In a popular article describing the process of an auditory brainstem implant, there appears this statement: "the area where the axons (nerve fibres) and cochlear nucleus (synapses)—which transport sounds picked up by the ear to the cerebral cortex—are found".[7] In this case, sound appears (in similar manner to the two ANSI definitions) to exist simultaneously in two states: the vibrations to be picked up by the ear and the auditory sensation of such vibrations. Such ambiguity, vagueness, and imprecision (perhaps surprising to those expecting the claimed-for exactitude, rigour, and precision of the natural sciences) can also be found in scholarly textbooks and academic articles on acoustics and psychoacoustics as Pasnau makes clear when he enumerates several examples that, taken together, bring into focus the incoherency of the standard physics-based definitions of sound. For instance, in the same 19th century acoustics textbook by John Tyndall, sound is located within the brain, within the sound source, or is to be found moving within the medium[8] while, in a modern handbook dealing with the perception and cognition of hearing, sound is described both as a compression wave and as the sound source itself.[9] Finally, an acoustics textbook I have long used for teaching (and from which I myself was taught) describes sound as "a mechanical disturbance of the medium".[10] While the authors admit that it is a simplistic definition, it is nevertheless the one provided on the opening page of this standard textbook and there are so many flaws with it that it is difficult to know where to begin.

5 American National Standard, *Acoustical Terminology*. ANSI/ASA S1.1-2013

6 This definition has the following note appended: "Not all sounds evoke an auditory sensation, e.g., ultrasound or infrasound. Not all auditory sensations are evoked by sound, e.g., tinnitus." From part b) of the definition, sound is defined as an auditory sensation; what the note is clarifying [sic], then, is: "Not all sounds evoke a sound [...] Not all sounds are evoked by sound." I suspect the reader may be as puzzled as I am by such claims while secretly applauding such pataphysical dadaism from the normally po-faced field of acoustics.

7 D. Pouliot, *Hearing Without Ears (Auditory Brainstem Implant)*. Accessed August 25, 2015, http://lobe.ca/en/non-classee/hearing-without-ears-auditory-brainstem-implant/#.VdQuvbfGam4.

8 On one page alone, the following statements occur: "It is the motion imparted to this, the *auditory nerve*, which, in the brain, is translated into sound" and, discussing exploding gases in a lecture theatre, "every ear in this room is conscious of a shock, to which the name of sound is given". On a later page, the following appear: using the analogy of balls in a row hitting against each other (thus the motion of the first ball is transferred to the last), Tyndall states that "thus is sound conveyed from particle to particle through the air" and yet, when describing how this motion sets the tympanic membrane vibrating, which motion is itself transmitted along the auditory nerve to the brain, it is in the brain that "the vibrations are translated into sound". Tyndall's book was written four years after, and is heavily indebted to, von Helmholtz's *Die Lehre von den Tonempfindungen*, a work of some significance in the canon of modern acoustics. J. Tyndall, On Sound, (London: Longmans, Green, and Co., 1867), 2-4.

9 R. Pasnau, "What is Sound?" The Philosophical Quarterly 49:196 (1999): 318-319.

10 D. M. Howard and J. Angus, Acoustics and Psychoacoustics, (Oxford: Focal Press, 1996), 1.

Problems with these definitions

As there are rather many definitions of sound as a physical phenomenon, to avoid confusion I shall group them together under what Pasnau terms the *standard view* of sound.[11] For Pasnau, this term brackets together two common concepts found within all these definitions. First, as Aristotle first identified, that sound is the object of hearing and, second, that sound exists in the medium (for humans, typically air) and thus sound in the standard view, however the definition is formed, is the object of our hearing and this object consists of vibrations or oscillations in a medium. The second definition provided by ANSI – that sound is an auditory sensation – complicates matters as it might be said to reference Aristotle's statement on sound but it does not directly reference the second common concept that Pasnau identifies. It is, however, best viewed not as a physics-based definition but as a psychology-based definition. Nevertheless, psychoacoustics both uses this definition and relies for it on the presence of sound waves.

There are two fundamental categories of problems with the quite singular standard view of sound. First is the issue of the language and the words used in the definitions bracketed under that view. Second is the quite limited usefulness of such definitions.

First, the wording. In the OED definition, which or whose ear does the sensing of the vibrations? Human ears only or those of other animals, birds, and insects? What precisely is an ear – is that of reptiles, such as snakes, properly termed an ear when it comprises only what may be described (in human terms) as the inner ear? What of the case of auditory brainstem implants where the useless ear (or auditory nerve) is bypassed with a microphone – do the recipients of such implants sense sound without the use of ears? What of tinnitus – particularly subjective forms such as the high-pitched ringing experienced by many – where are the vibrations or oscillations in a medium?

If such tinnitus sufferers do indeed hear sound, then the definitions provided by ANSI can be discarded out of hand. Even my attempt to use jargon to more accurately define sound while remaining true to the standard view fails when one asks two simple questions: what precisely is the auditory system and what is the frequency range of hearing?[12] Across hearing species, there are widely varying hearing frequency ranges. If sound is indeed a pressure wave/vibrations/oscillations in a medium that can be sensed by the ear/auditory system, then, as a thought experiment, if both a human and a dog are subject at the same time to the same sound wave at 30kHz, why should it only be defined as sound for the one simply because the other's auditory sensory apparatus cannot sense it? And what of infrasound that can be felt through the skin – is the auditory system simply part of our tactile sense because, as one famous (and famously profoundly deaf) musician has suggested, "hearing is basically a specialized form of touch"?[13]

Related to this problem of the wording is the use made of the definitions because the use that sound as a definition is put to depends to a large extent on the firmness (or not) of the wording (although definitions can be blithely disregarded as is the

11 Pasnau. "What is Sound?" 309-310.

12 A good human hearing range is typically quoted as between 20Hz and 20kHz.

13 E. Glennie, *Hearing Essay*, (1993). Accessed August 19, 2015, https://www.evelyn.co.uk/hearing-essay/.

case previously noted with Tyndall happily defining sound as a motion propagated through air while at the same time stating that sound exists in the brain). If one cannot agree on one rigorous definition of sound, then how does one know precisely what the object of study is (let alone hearing) or, indeed, if two articles purporting to deal with sound are, in fact, dealing with that same object of study or hearing?

The second category of problems is, for me, the more egregious and the reason why I have spent some time devising what I believe to be a more precise, comprehensive, and useful definition of sound: the standard view of sound is too limited in its scope; it does not match the reality of our experience and thus is too limited in its use-value. My objection is founded upon two disparate frameworks of knowledge that might broadly be defined as objective and subjective. The physics-based definition attempts to be objective because industrialization requires standardization and mass production and so two people subject to the same sound wave at the same time and at the same location will (must), according to such definitions, hear the same sound. Similarly, assuming all things technological are equal, the same audio data played back at different times will (must) produce the same sound waves particularly if that audio data is in digital format. This, at least, is the theory; endlessly reproducible sonic artefacts, identical copies of some original, and all neatly packaged for the digital consumer.

The reality of our experience is somewhat and frequently different and this, I believe, is where the standard view of sound fails. Two simple and easily accessible demonstrations will suffice to show this. The first is the well-known McGurk Effect (numerous examples of which can be found on the Internet including one produced by the BBC[14]). Briefly, the syllable 'baa' is spoken, recorded, and played back over two videos, one in which the subject articulates 'baa' (from whence the syllable 'baa' is recorded) and the other in which the subject articulates 'faa' (the lip and tongue movements are quite distinct between the two). Thus, the same audio data is reproduced multiple times and so, on the same equipment and settings and in the same location, the same sound waves move through the air and so, according to the standard view of sound, the same sound is reproduced time and time again. If sound is the sound wave then the same sound should be heard at all times by all listeners or even the single listener. This does not happen. Instead, what is heard are the syllables 'baa' and 'faa' over the appropriate image as if two different sets of sound waves are being produced. If part of the standard view of sound is that sound is the object of our hearing, it is interesting to note that listeners, when asked what they hear, report the two different sounds (i.e. the objects that they hear) despite the presence of only the one recorded sound wave. One must conclude either, that if sound really is a sound wave, then sound is not the object of our hearing or, that if sound really is the object of our hearing, then sound waves are not sound. Either way, the standard view of sound is as incoherent as Pasnau claims.

On the issue of the effects of cross-modality of hearing and vision (and other senses) and how that issue affects our understanding of what sound is, I return to

14 http://www.youtube.com/watch?v=G-lN8vWm3m0

this later. The second demonstration is what might be broadly termed mimicry if not outright trickery and it trades on ambiguity of the source and/or inexperience of the listener but, unlike the McGurk Effect, does not use other sensory modalities. There are two examples I present here both of which are easily testable. The first is to play for a subject the recording of a Geiger counter. Anyone familiar with the characteristic crackling sound will invariably give the correct answer upon being asked what the sound is but others will give different answers such as the preparation of popcorn. My favourite example, though, is the extraordinary mimicry of the Australian lyrebird that has a repertoire of sounds including camera shutter (replete with automatic winding motor), chainsaws, alarms, and sirens. Those not familiar with this creature and its ability to mimic such sounds will typically label the sounds as listed above (assuming familiarity with these objects) while those familiar with the widely available examples of the bird's mimicry on the Internet will immediately label the sounds as the lyrebird.

In the above paragraph, the astute reader will have noticed a subtle shift in my use of language. It is a shift in the use of the word 'sound' that now equates it with the objects that (originally) produced the sound waves that were recorded for these two examples. This relates to our everyday experience of sound that, again, I return to later being content for now to state that this everyday definition of sound is not one that finds favour with the standard view of sound (in this everyday definition, sound is the sound wave *source* – logically, the sound produces the sound wave and so sound waves supervene upon sounds – sound waves, then, are how sounds are made manifest[15]).[16]

Other definitions of sound

It is not my intention to thoroughly detail or even list all definitions of sound both historical and current. Here, I present definitions that, in many respects, oppose the standard view of sound although some still rely on the presence of sound waves (which, in these definitions are not themselves sound); the definition encapsulated within *sonic virtuality* that I present later does not rely on this presence. For the reader wishing to peruse a fuller list of sound definitions, I refer them to Nudds and O'Callaghan and especially to Casati and Dokic.[17] [18] Those I present here are, I believe,

15 R. Casati and J. Dokic, "Some Varieties of Spatial Hearing," in Sounds & Perception, eds. M. Nudds and C. O'Callaghan, (Oxford: Oxford University Press, 2009), 98-99.

16 Even those adhering to the standard view of sound fall prey to this equivalence between sound and sound source as Pasnau describes elsewhere in this article and as in this example from an acoustics textbook: "Trained singers [...] exhibit an effect known as 'vibrato' in which their fundamental frequency is varied [...]" (Howard and Angus. Acoustics and Psychoacoustics. 193). Thus, the fundamental is a property of the singer and not a vibration in the medium of air. (I should state here that, while it can be fun, it is not my intention to attack such trivial and venial deviations from scientific orthodoxy – I would quickly become a hostage to fortune myself were I to do so – but rather to demonstrate the yawning chasm and barely sustainable tension that exists between the standard view of sound and our everyday experience and understanding of sound.

17 M. Nudds and C. O'Callaghan, eds., Sounds & Perception, (Oxford: Oxford University Press, 2009).

18 R. Casati and J. Dokic, "Sounds," Stanford Encyclopedia of Philosophy (2005/2010). Accessed June 3, 2014, http://plato.stanford.edu/entries/sounds/.

useful in this article's context because they begin to take on a phenomenological flavour where the focus shifts from the physical properties of, and changes in, a medium toward a focus upon sensation, perception, and cognition. Furthermore, those definitions that I do describe below are reasonably current as of date of writing. One should not, however, make the mistake of assuming that stressing the importance of perception to any study of sound or, indeed, defining sound in perceptual terms is anything new. Bregman's work, for example, and although one must assume that he holds to the standard view of sound because "sound enters the ear",[19] is important in that it is an earlier statement in favour of analysing the perception of sound rather than sound wave properties and even earlier instances of this type of thinking (although dealing with music rather than sound) can be found in Carl Stump's *Tonpsychologie* – an early example of phenomenology.[20]

Before I begin, though, I wish to briefly present another physics-based definition of sound that is of some interest and this is the conception of sound as comprising particles. This is an ancient idea that has recently (in its bare bones conception at least) been given a new lease of life. The Greek philosopher Democritus, an early proponent of classical atomism active during the 5th and 4th centuries BC, proposed that sound was a stream of atoms emitted by a thing. What is of interest here is the description of sound as comprising particles (the atoms) that was (re)proposed in 1947 by Gabor;[21] a concept that has found use both in the fields of granular synthesis and quantum physics (where the sound particle is termed the *phonon*[22]).

Sound as a property of an object

Pasnau, in his argument for the incoherency of the standard view of sound, suggests instead that sounds reside "within the object that 'makes' them [...] sounds are either vibrations of such objects, or supervene on those vibrations".[23] While this statement is ambiguous (sound is either a vibration or is contingent upon a vibration), Pasnau is quite clear about the location of the sound – it is within the object and thus "objects *have* sounds".[24] This question of the location of sound, rightly so, is fundamental to the genesis of any definition of sound, no less for Pasnau who appeals to the veridicality of our auditory perception. This perception informs us that sound, like colour, is located at the object and is part of the external world (hence we can use the location of sound to locate objects in that external world): "we do not hear sounds as being in the air; we hear them as being at the place where they are generated".[25]

19 A. S. Bregman, Auditory Scene Analysis: The Perceptual Organization of Sound. (Cambridge, Massachusetts: MIT Press, 1990), 7.

20 C. Stumpf, Tonpsychologie. (Leipzig: Verlag Von S. Hirzel, 1883).

21 D. Gabor, "Acoustical quanta and the theory of hearing," Nature 159 (1947): 591–594.

22 Specifically, it is the long-wavelength phonon (the acoustic phonon – the curious concept of the 'sounding sound' or 'sound having to do with sound' if one were to translate this literally) that gives rise to sound (thus, the phonon itself is not sound).

23 Pasnau. What is Sound? 316.

24 Ibid. 316.

25 Ibid. 311.

To locate sound elsewhere is a denial of the veridicality of our perception, a denial that, for Pasnau, is a step too far as it "shakes the very foundations of our cognitive assumptions".[26]

Sound as an event

Basing his idea upon Aristotle's statement that "sound is a particular movement in air"[27] – a statement that *prima facie* appears to support the standard view of sound – O'Callaghan argues that sound is an event that sets a "medium into periodic motion".[28] He arrives at this definition via two other passages in *De Anima*: "everything which makes a sound does so because something strikes something else in something else again, and this last is air" and "sound is the movement of that which can be moved".[29] In this definition, sound waves, and indeed our perception of them, are mere by-products of sound events and these events, comprising as they do movement, involve change and therefore take place over time.

Sound as both object and event

Here, I come to the first phenomenological definition of sound in which sound neither possesses physicality nor is it the property of a physical (and external) object. For Scruton, sound is both object and event; that is to say, sounds are secondary objects and pure events because they do not undergo change, as physical objects do, and they do not happen to anything.[30] Scruton's approach is closer to the valid, veridical definition of sound that I aim for because it is founded upon an assessment of "those features of sound that make sound so important to us [...] socially, morally, and aesthetically".[31] Furthermore, Scruton takes account of acousmatic phenomena arguing that sound (e.g. in the case of radio broadcasts or audio recordings) suffers no loss in coherence when divorced from its physical source and no loss in what "is essential to the sound as an object of attention".[32] This perceptually based definition of sound is heavily influenced by studies in music perception including those phenomena that produce "a virtual causality that has nothing to do with the process whereby sounds are produced";[33] the streaming and grouping of musical notes based upon their similarity in timbre or pitch and their temporal proximity to each other.

26 Ibid. 316. That Pasnau uses the word 'assumptions' does at least, in my view, keep the door ajar for such a denial.
27 Quoted in C. O'Callaghan, "Sounds and Events," in Sounds & Perception, eds. M. Nudds and C. O'Callaghan, (Oxford: Oxford University Press, 2009), 27.
28 Ibid. 37.
29 Quoted in Ibid. 27.
30 R. Scruton, "Sounds as Secondary Objects and Pure Events," in Sounds & Perception, eds. M. Nudds and C. O'Callaghan, (Oxford: Oxford University Press, 2009), 50.
31 Ibid. 62.
32 Ibid. 58.
33 Ibid. 64.

Non-cochlear sound

Another phenomenological definition of sound is presented by Riddoch who states that sounds "first and foremost [are] worldly phenomena".[34] For Riddoch, the standard view of sound has failed in "demonstrating a causal mechanism linking our neurological processes with the supposed subjective effect – the world of our perception"[35] and, furthermore, that view does not take into account the idea that sounds "are always in the first instance meaningful sounds".[36] On this basis, he argues that cochlear sound (i.e. that sound involving sound waves) is in fact non-cochlear sound, a group in which Riddoch also includes synaesthetic sounds (see cross-modality below), infrasonic sounds that are detectable by other body parts than the ear (see, for example, Glennie[37]), and auditory imagination.

Everyday listening

While he does not provide a definition of sound per se, in his focus on what he terms everyday sounds and everyday listening, Gaver also objects to the standard view of sound and, for this reason, his work merits inclusion here.[38] [39] Importantly, Gaver's everyday sounds are non-musical sounds. He therefore focuses on sounds that are inharmonic and noisy because these, he explains, have never been the subject of acoustics or psychoacoustics, two fields that prefer to concentrate on periodic and therefore pitched sounds.[40] While this situation has changed somewhat in the intervening years, Gaver's complaint that studies conducted under the aegis of the standard view of sound display methodological bias still holds. This can be summed up neatly in a statement provided by Gibson (by whom Gaver was influenced): the standard view of sound "treats physical sound as a phenomenon *sui generis*, instead of as a phenomenon that specifies the course of an ecological event; sound as pure physics, instead of sound as potential stimulus information".[41] Thus, for Gaver, an incomplete and inaccurate view of sound is formed when a) only musical tones are studied, and the results then used to build up a picture of sound and auditory perception and b) the language used to describe sound (frequency, amplitude, and so on) is inadequate to the task of describing our everyday experience of sound in which we typically equate sound and sound wave source – a subject is played a sound and asked by the white-coated psychologist to describe what has been heard; 'I hear a plane', the subject responds much to the frustration of the psychologist who requires the answer

34 M. Riddoch, "On the Non-cochlearity of the Sounds Themselves" (paper presented at *International Computer Music Conference*, September 9-14, 2012, 14).

35 Ibid. 14.

36 Ibid. 13.

37 Glennie. *Hearing Essay*.

38 W. W. Gaver, "What in the World do we Hear? An Ecological Approach to Auditory Perception," Ecological Psychology 5:1 (1993): 1–29.

39 W. W. Gaver, "How do we Hear in the World? Explorations in Ecological Acoustics," Ecological Psychology 5:4 (1993): 285–313.

40 Gaver. How do we hear in the World? 286-287.

41 J. J. Gibson, The Senses Considered as Perceptual Systems, (Boston: Houghton Mifflin, 1966), 86.

to be couched in the scientific language and terminology seen as appropriate to the study of sound.[42]

Problems with these definitions

There are a number of issues that can be raised from the above definitions and that leave me unconvinced about each of them. The first I wish to deal with is the location of sound whether that sound is an object, a property of an object, or an event. Equally, the standard view of sound displays some incoherency within it as to the location of sound, if not in its definition(s) (although the two ANSI definitions given above provide two distinct locations) then in the lax ways to which these definitions have been put. (These I have noted above.)

The location of sound is always given in relation to the organism that hears the sound (thus, from the point of view of location of sound, sound is always the object of hearing whatever its definition). The locations may therefore be roughly grouped as distal, medial, or proximal in relation to the listener. The standard view of sound (and the first ANSI definition) encapsulates the view that sound location is medial; it is located, as a vibration, in a medium between the sound source and the listener. This begs a number of questions not least of which is *is sound therefore static or mobile?* Here, we run into further problems with the standard view in that sound (in such definitions) does indeed move through the medium but it surely must initially be distal (as it issues forth from the sound source not the medium) and ultimately proximal (this, presumably, is where the second ANSI definition takes over – sound is auditory sensation originating in the ear of the listener).

Our everyday experience of the location of sound is typically that it is distal. That is, the sound is located at the sound source, hence our everyday equation between sound and sound source. This experience is expressed in terms such as 'the knocking is inside that wardrobe.' But we also experience sounds as mobile (while the sound source is static) as in 'the knocking is coming from that wardrobe.' Equally, we can also experience sound inside the head and this need not be tinnitus or auditory hallucination or the imagination of sounds as those who have experienced in-head localization of sound through headphone use will be familiar with. One should therefore be cautious with and somewhat mistrustful of definitions of sound (or at least theories about the location of sound in which location is fundamental to such definitions of sound) that appeal to everyday experience as there is an incoherence here too, no less than in the standard view of sound.

Thus, one should be distrusting of the sound as the property of an object definition provided by Pasnau because it is based on the supposed everyday location of sound as being at the sound source (the object). Equally, there is an incoherence in Pasnau's view when the case of cinema (see further below) is used; sound is located in our everyday listening on the screen yet there are no vibrations there of the type that Pasnau claims are the sounds themselves. Regarding sound as an event, O'Callaghan is

42 Gaver. How do we Hear in the World? 286-287.

frustratingly vague in stating that "sounds are events that take place near their sources, not in the intervening space"[43] while being explicit that "sounds are stationary relative to their sources".[44]

On the location and locatedness of sound, the evidence from neuroscience is of interest as are the observed effects of cross-modality, particularly those effects arising from the interoperation of vision and hearing. In the first instance, I specifically refer to neuroplasticity as it relates to the auditory system. In auditory neuroscience, the localization of sound refers to the ability to locate the sound wave source *viz.* the source of the sound. Thus, in this branch of science, sound is located distally – to locate the sound is to locate the source of the sound because the sound is located at that object or event. The book on which this article is based is able to go into far greater detail on the subject of neuroplasticity than I have the space to do so here, therefore I will limit myself to some brief notes on the most pertinent points to be made.

There are a number of studies (on humans, other animals, and birds) demonstrating that the artificial blocking of one ear displaces (in some cases significantly) the horizontal localization of sound when compared to localization using two unimpaired ears (binaural hearing).[45] (See, for instance, Slattery and Middlebrooks or Kacelnik and colleagues.[46] [47]) This phenomenon is as expected if one subscribes to the Duplex Theory of sound localization.[48] By itself, this raises questions as to the location and act of locating sound – one could infer, for example, that if sounds are sound waves (which radiate out from a source) then sounds can be located as being in or from a direction that is different to where the sound wave source is if binaural hearing is impaired (the same effect is not noticed with vision if one eye is shut). In this case, and logically, sound is not located at the sound source but is mobile and in the medium. Nevertheless, these neurological studies prefer to state that sound is being *incorrectly* localized (because sound is distal and located at the sound source in this science).

43 O'Callaghan. Sounds and Events. 48. This intervening space is the medium between sound source and listener and through which the sound waves (that are the effects of sounds) travel. What the space or locus is between the sound source and the medium in which the event takes place, O'Callaghan never states. One should also ask: precisely how near is 'near'?

44 Ibid. 46.

45 See, for example, F. L. Wightman and D. J. Kistler, D (1997). "*Monaural Sound Localization Revisited,*" Journal of the Acoustical Society of America 101: 2 (1997): 1050–1063 for some objections to these studies on the basis that full monaural hearing is never achieved and that the test signals used do not include everyday sounds.

46 W. H. Slattery III and J. C. Middlebrooks, "Monaural Sound Localization: Acute Versus Chronic Unilateral Impairment," Hearing Research 75:1-2 (1994): 38–46.

47 O. Kacelnik, F. R. Nodal, K. H. Parsons, and A. J. King, "Training-induced Plasticity of Auditory Localization in Adult Mammals," PLoS Biology 4:4 (2006): 0627–0638.

48 Briefly, organisms with binaural hearing use two overlapping channels of auditory information in order to localize sound; for humans, lateral sound sources can be localized using *interaural level difference* if frequencies in the sound wave are above about 640Hz and *interaural time difference* if frequencies in the sound wave are below about 740Hz. Ambiguity about the direction of sound sources directly ahead or behind and for sound source position on the vertical plane are resolved through assessing the spectral filtering effects of pinnae and through the simple expedient of moving our heads.

Some of these studies and others demonstrate that such mislocalization of the sound source (although it remains accurate localization of the sound wave or some point along its front) can be corrected over time and such re-localization facilitated and hastened along with training.[49] The brain's processes and physical structures for localization change and adapt to the new monaural conditions and such neuroplastic ability is maintained throughout life.[50] Interestingly, some studies suggest that the brain can store more than one of these auditory localization schema, switching almost immediately among them as the subject's monaural and binaural conditions change.[51]

Outside of neuroscientific studies, there are numerous everyday examples of the human ability to localize the sound in some direction or at some spot other than from where the sound wave travels or the sound wave source is. The prime example of this is what is known in psychoacoustics as the ventriloquism effect[52] or in film studies and related areas, with some minor differences to the ventriloquism effect, as synchrony or synchresis.[53] [54] Briefly, synchresis is the cross-modal effect of multi-modal stimuli (in this case, visual and aural) being perceived as one event and it is an effect that is fundamental to that cinema that uses diegetic sound. It relies on the ability to separate sound from the sound source (sound waves and the cinema's loudspeakers) and to localize the sound not at the sound source but on events, objects, or characters depicted on the screen. This bears some relation to the McGurk Effect discussed above in that the sound heard depends to some extent on the image seen but the difference is to do with the location and localizing of sound. As with the neuroscientific evidence from sound localization studies above, synchresis is a process that relies on the ability to locate sounds other than where their sound wave sources are and, thus, to perceive that the sound travels in a direction other than the direction the sound waves themselves radiate from.[55] This not only casts further doubt on the Duplex Theory, but also further reinforces my doubts about the standard view of sound and other definitions such as sound as property of an object or sound as event (located at or near its source).

49 e.g. P. M. Hofman, J. G. A. Van Riswick, and A. J. Van Opstal, "Relearning Sound Localization with New Ears," Nature Neuroscience 1:5 (1998): 417–421.

50 See, for example, M. Schnupp, I. Nelkin, and A. King, Auditory Neuroscience: Making Sense of Sound (Cambridge, Massachusetts: MIT Press, 2010).

51 e.g. Hofman et al. Relearning Sound Localization with New Ears.

52 See, for example, D. H. Warren, R. B. Welch, and T. J. McCarthy, "The Role of Visual-auditory "Compellingness" in the Ventriloquism Effect: Implications for Transitivity Among the Spatial Senses," Perception & Psychophysics 30:6 (1981): 557–564.

53 J. D. Anderson, The Reality of Illusion: An Ecological Approach to Cognitive Film Theory, (Carbondale and Edwardsville: Southern Illinois University Press, 1996).

54 M. Chion, Audio-vision: Sound on Screen, trans. C. Gorbman, (New York: Columbia University Press, 1994).

55 It may be that this is also an instance of neuroplasticity and an example of the ability to store and recall instantaneously multiple auditory localization schema. The cinema, of course, is not a laboratory scenario where binaural hearing is artificially impaired but I am unaware of studies conducted on cinema audiences looking for evidence of auditory neuroplasticity. One schema is learned and used within the artificial environment of the cinema or when watching television or playing computer games, with at least one other schema for use elsewhere where sounds ought be perceived as originating from the sound wave source – very important when crossing the road, for example. However, see below for an argument concerning our active locating of sound.

One other objection I have to some of the non-standard view definitions relates to the methodology used. This is particularly the case with Scruton's definition. My objection here relates to Scruton's insistence that humans alone have the ability to perceive order in sound (and thus they alone can perceive music as opposed to sound) and the basis of his thinking behind the definition being studies of music perception.[56] In the first case, this claim rests upon increasingly shaky ground.[57] Second, in limiting his thinking to music studies alone, Scruton derives a definition that is founded upon the wrong object of attention; music is neither sound alone nor is it alone sound. O'Callaghan displays a similar bias towards musical or pitched sounds (which I believe weakens his argument for a definition of *sound*) when he states that a sound event sets a "medium into periodic motion"[58] – not all sound waves are periodic and thus O'Callaghan's definition is not for all sounds.

The intention of this section has been to briefly enumerate and examine a number of questions concerning the standard view of sound and other definitions of sound. This is because, to my thinking, there are inconsistencies in the definitions and incoherency within some of the more accepted definitions and this leads me to question their validity and therefore their effectiveness and suitability of purpose.[59] The next section presents an alternative definition. As with the definitions and their problems listed above, the presentation of this new definition is necessarily brief. Its full exposition can be found within the book on which this article is based.

Sonic Virtuality

> *Sound is an emergent perception arising primarily in the auditory cortex and that is formed through spatio-temporal processes in an embodied system*

This definition of sound is the one that my co-author Tom Garner and I developed first in response to our many objections to other definitions that we deemed wholly unsatisfactory and/or incoherent and, second, as a result of several years collaboration particularly in the area of biofeedback and computer-game audio. There is no intention to replace the standard view of sound – briefly, that sound is a sound wave – although we contend that what is studied in physics and acoustics is not sound but sound waves; seen from that point of view, the standard view of sound (waves) is perfectly valid (if a little inconsistent in its application), verifiable, and usable (usability is what definitions are about, after all). Our intention is to provide a definition that a)

56 Scruton. Sounds as Secondary Objects and Pure Events. 62-66.

57 See A. D. Patel, J. R. Iversen, M. R. Bregman, and I. Shulz, "Studying Synchronization to a Musical Beat in Nonhuman Animals," The Neurosciences and Music III—Disorders and Plasticity: Annals of the New York Academy of Sciences (2009), 459–469 for an overview of claims for music perception abilities among non-humans.

58 O'Callaghan. Sounds and Events. 37.

59 I am not the first to point out the many inconsistencies and outright incoherency of the standard physics-based definition(s) of sound and here I only provide a few of the many objections that could be and are raised. Others providing fuller accounts include Pasnau (1999) and Riddoch (2012).

answers the many questions and supports and explains the positions we have about sound that are not answered or evidenced by other definitions and b) has a use-value particularly in the area in which we work, work that may be described as human-centred sound design. As stated, Tom Garner and I have worked and continue to work in the area of computer game audio and are especially interested in the real-time recognition of player affective states, feeding data on those states into the game engine, and using them to synthesize or process audio in response (which then re-engages with the players' affective states thus creating a biofeedback loop). This can be done and we have done it.[60] [61] What is difficult, though, is the recognition of emotions and affect beyond basic valence and arousal and what is lacking is comprehensive knowledge on the effects that sounds (rather than music) and changes in audio parameters in various contexts have on the individual's arousal state and emotions. Clearly, a model of sound founded upon a definition of sound that acknowledges the human factor, and particularly its subjectivity, will be of some use here. And this not only has use in my particular field but it also has use within other fields such as the field of recording music and sound. Before I move onto some thoughts about such usage, though, I need to unpack the definition above and to explain its genesis.

Sound is an emergent perception . . .
This is the core of our definition; sound is a perception within our minds – thus, the origin of sound is there – and its creation is an on-going, emergent process.

. . . arising primarily in the auditory cortex . . .
This emergent perception is initiated in a corporeal system and is centred upon the auditory cortex. A number of factors together contribute to the emergence of sound as these include both corporeally external factors (the *exosonus*) such as sound waves and stimuli across other sensory modalities, and internal factors (the *endosonus*) such as other perceptions (thus cross-modality), emotions, memory, reasoning, knowledge, and so forth.

. . . and that is formed through spatio-temporal processes . . .
The emergence of sound takes place over time and is inherently spatial being part of the process of locating self in the world.

. . . in an embodied system.
Following ideas expressed in the field of embodied cognition, the perception that is sound is a brain- and body-based function and is indivisible from the wider environment – the embodied system in which sound emerges comprises brain, body, and environment.

60 M. Grimshaw and T. Garner, "Embodied virtual acoustic ecologies of computer games," in The Oxford Handbook of Interactive Audio, eds. K. E. Collins, B. Kapralos, and H. Tessler, 181–195. (New York: Oxford University Press, 2014).

61 T. A. Garner and M. Grimshaw, "The Physiology of Fear and Sound: Working with Biometrics toward Automated Emotion Recognition in Adaptive Gaming Systems," IADIS International Journal on WWW/Internet 11:2 (2013).

From this definition, several key propositions and assertions arise that are more fully stated and supported in the book:

- It is the sonic aggregate that is the engine for the emergence of sound and it comprises fluid and dynamic spatio-temporal relationships of sensuous/nonsensuous, material/immaterial things and phenomena
- The set of material, sensuous components of the sonic aggregate we term the exosonus – it is *not* a requirement for the emergence of sound
- The set of immaterial, nonsensuous components of the sonic aggregate we term the endosonus – it *is* a requirement for the emergence of sound
- Using Deleuzian terminology[62], all components of the sonic aggregate, whether exosonic or endosonic, are actual as is sound. The actualization that is the emergence of sound as a perception derives from processes that are themselves virtual
- It is the force of potential in the fluid, dynamic sonic aggregate that, when achieved under particular spatio-temporal conditions, leads to the actualization itself
- *The hear and now* of sound is the emergent actualization in the here and now and it requires the presence of a perceiver
- In order to locate ourselves, we cognitively offload the location of sound onto the environment – this is an active (re)location of sound
- Sound is meaningful because different types of meaning form part of the endosonus; sound waves are inherently meaningless
- A sound wave may be acousmatic but a sound is never acousmatic. Hence, reduced listening (the concept of a sound divorced from source, cause, environment, and meaning – and thus listenable to in and of itself) remains a theoretical concept not achievable in practice
- Discussions of real sound and virtual sound are invalid as there is no such distinction
- The emergence of sound can determine epistemic perspective as epistemic perspective can determine the quality and nature of the emergent sound
- Imagined sound is as much sound as that sound perceived in the presence of sound waves.

It should be immediately clear from the above that our description is, in large part, founded upon the Deleuzian concept of the virtual. Thus, the sonic aggregate has some equivalence to the concept of the virtual cloud. It is from this sonic aggregate, with its dynamic tension and shifting potentials, that sound is actualized as an emergent perception. This aggregate comprises the endosonus and may optionally, as it frequently does, also comprise exosonic components such as sound waves. In the assertion that sound waves are not necessary to the emergence of the sound perception, the definition thus accounts for a wider and, to my mind, more valid view of what sound is than the standard view is capable of doing. In doing so, it tackles a number of in-

62 e.g. virtuality encompasses the idea of a dynamic *virtual cloud* of potentiality from which *actualization* occurs. It is important to note that, in this conception of virtuality, the virtual is a mode of the real and not to be contrasted to the real (*pace* real worlds and virtual worlds).

consistencies and sonic anomalies at odds with this standard view such as tinnitus (which is sound) and how the McGurk Effect is perceived as different sounds despite the presence of the same sound wave.

I have previously stated that the question of the location of sound is fundamental to many definitions of sound; if one can decide *where* sound is then one is halfway to defining *what* sound is because the location is descriptive of the sound's environment. I do not think that the concept of sonic virtuality is likewise so dependent upon the location of sound but, nevertheless, I feel bound to use the definition and concept to explain the location of sound.

How is it that, for something so fundamental to the daily living of so many organisms including humans, for something that for so long has been an object of intense study, and for something for which the standard view of sound offers a medium-specific location, there is so much disagreement about the location of sound? Perhaps there is something wrong with all those definitions. The concept of sonic virtuality offers this explanation: the origin of sound is in the secondary auditory cortex[63] but, where the mind is a system comprising brain, body, and environment, sound can be located anywhere in this system. This is not to say that the definitions founded upon or arguing that sound is distal, medial, or proximal are all, and at one and the same time, correct. Far from it. Such definitions use the passive form of 'to locate' and so sound is to be found somewhere, typically out there in the world.

Sonic virtuality, on the other hand, uses 'to locate' in the active sense. We ourselves locate, that is, place (rather than find) sounds somewhere within the system of which the mind is comprised. This is the cognitive off-loading that is one of the tenets of embodied cognition. This locating of sound is learned and develops through early childhood into an automatic act of placement and it explains the close and cross-modal relationship to vision as we, while babies, cast around, looking for the source of a sound wave, beginning to learn the associations between movement, size, and material and sound waves. It is a skill that can be lost and relearned (as neuroplasticity demonstrates) and it can be adjusted at will to new physical environments (such as the cinema or when playing a computer game where the location of a sound wave in air or the location of the origin of that sound wave are very different to those places where we actively locate the sound). The ability to actively locate sound arises from the necessity of being and acting within an ecology and to listen is to be present in a world, whether virtual or real – sound is the means not to locate objects in an external world but rather to locate our selves in relation to other objects.

Clearly, and as I have stressed throughout, an article cannot go into the detail that a book can; *Sonic Virtuality* is a more substantial and substantiated treatment of the subject matter and, in addition to what has been cursorily covered here, also includes evidence and support for the concept from the virtual acoustic environments of com-

63 This is justified more substantially in the book but, briefly, the assertion arises because this is the common locus where activity is observed when sound is perceived either in the presence of sound waves or through an act of imagination, bidden or unbidden.

puter games, the philosophy of knowledge and belief, auditory imagery, hallucination and other pathologies, imagination, and creativity and is at pains to demonstrate the practical purpose of the definition. It is on this last point that I will close the article. The presentation upon which this article is based was delivered at a conference on sound and music production and so, although the book gives different examples, I will provide an argument for the use-value of the definition and concept of sonic virtuality within the domain of sound design and will do this by means of a brief example of the concept as it might be used.[64]

If, as our definition asserts, sound is a perception that arises in the brain, then the obvious question to ask is *can we extract it*? Before I look more closely at this question, I need to clarify the terminology, particularly the word 'extract.' If sound is a perception in the brain, then, once extracted from the brain, it is no longer sound. To be more precise, then, 'extraction' is used metonymically here to stand for the transposition of sound to audio data. My speculative answer to the question then is, and for reasons outlined below, *no, not yet, but it will be possible in the near future*. Thus, one should begin to lay the groundwork for this possibility and the necessity of this is one of the reasons for devising the concept of sonic virtuality.

Neural decoding, as it is known within neuroscience circles, is a developing field used, among other purposes, for the extraction of still and moving images from the visual cortex.[65] Neural activity is decoded and visually represented on a monitor and this decoding is accurate enough to be able to recognize features of the image the subject is concurrently gazing at. There is work afoot on extracting speech[66] and plans to extract music from the brain[67] but, first, there are good reasons to state that music and speech are conceptually and perceptually distinct to sound (even if somewhat and sometimes related) and, second, there are different processes and different regions in the brain involved.[68] [69] Previous and current composers and performers that make

64 One possibility that did not make it into the book concerns the cross-modality of sound and smell perceptions and its relevance to presence in virtual environments. See Grimshaw, M. and Walther-Hansen, M., "The Sound of the Smell of My Shoes," (in proceedings of the *10th Audio Mostly Conference*, October 7-9, 2015).

65 See, for example, S. Nishimoto, A. T. Vu, T. Naselaris, Y. Benjamini, B. Yu, and J. L. Gallant, "Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies," Current Biology 21 (2011): 1641-1646. I must stress here that the orthodoxy of neuroscience does not express the same concept of image as I do of sound. That is, in this case, vision is the perception of an image and that perception is not the creation or emergence of an image itself. Thus, what is being extracted is not the image but the perception of that image.

66 B. N. Pasley, S. V. David, N. Mesgarani, A. Flinker, S. A. Shamma, N.E. Crone, et al., "Reconstructing Speech from Human Auditory Cortex," PLoS Biology 10:1 (2012).

67 J. Thompson, M. Casey, and L. Torresani, "Audio Stimulus Reconstruction using Multi- source Semantic Embedding, in Neural Information Processing Systems," (paper presented at the *Neural Information Processing Systems Workshop on Machine Learning and Interpretation in Neuroimaging*, December 9-10, 2013).

68 R. J. Zatorre, P. Belin, and V. B. Penhune, "Structure and Function of Auditory Cortex: Music and Speech," *Trends in Cognitive Sciences* 6:1 (2002): 37-46.

69 R. J. Zatorre and A. R. Halpern, "Mental Concerts: Musical Imagery and Auditory Cortex," *Neuron* 47 (2005): 9-12.

use of data gathered from brain activity are merely using audification and sonification techniques[70] rather than the extraction of sound (or music or speech).

The definition of sound presented here, and especially the concept of the sonic aggregate, is necessary to any thinking on how to extract sound.[71] It makes sense to conceive of sound as emerging from something like a sonic aggregate in the context of phenomena such as the McGurk Effect and when aware of the effect of higher order cognitive processing on auditory perception[72] and the effect of cross-modal sensation on auditory perception.[73] [74] [75] Sound may arise primarily in the auditory cortex but there are other sensory, perceptual, and cognitive factors involved too and these must be taken into account when attempting to extract sound.

Let us assume that, sometime in the future, it becomes possible to extract sound from the brain. Then what? If the extraction of sound is the transposition of sound to audio data then it becomes possible to conceive of, and design the possibility for, a process of sound design whereby the sound designer imagines and projects sound into a digital audio workstation. Thus, the sound conjured up in the absence of sound waves becomes available for manipulation as audio data, for copying, and for reproduction in the form of sound waves. This will fundamentally change the role of sound designer. Naturally, as sound in our concept is highly individual, the sounds emerging from the sonic aggregates of others then exposed to these sound waves are likely to have some differences to the original. This exemplifies the inherent subjectivity of the emergent perception and the necessity for the involvement of the perceiver that becomes evident from any study of sound – the sound is the object of perception, the sound wave is the object of sensation – and this difference is what missing from the standard view of sound and neither is it fully accounted for by other definitions of sound.

70 See M. Ortiz, "A Brief History of Biosignal-Driven Art: From Biofeedback to Biophysical Performance," *econtact* 14:2 (2012) for an overview.

71 Or a definition and concept that are similar.

72 e.g. N. Bunzeck, T. Wuestenberg, K. Lutz, H-J. Heinze, and L. Jancke, "Scanning Silence: Mental Imagery of Complex Sounds," NeuroImage 26:4 (2005): 1119–1127.

73 M. Hoshiyama, A. Gunji, R. Kakigi, "Hearing the Sound of Silence: A Magnetoencephalographic Study," NeuroReport 12:6 (2001): 1097–1102.

74 J. Voisin, A. Bidet-Caulet, O. Bertrand, and P. Fonlupt, "Listening in Silence Activates Auditory Areas: A Functional Magnetic Resonance Imaging Study," The Journal of Neuroscience 26:1 (2006): 273–278.

75 R. J. Zatorre, "There's More to Auditory Cortex than Meets the Ear," Hearing Research 229 (2007): 24-30.